On the Diversity-Multiplexing Tradeoff in Multiple-Relay Network

Shahab Oveis Gharan, Alireza Bayesteh, and Amir K. Khandani
Coding & Signal Transmission Laboratory
Department of Electrical & Computer Engineering
University of Waterloo
Waterloo, ON, N2L 3G1
shahab, alireza, khandani@cst.uwaterloo.ca

Abstract

This paper studies the setup of a multiple-relay network in which K half-duplex multiple-antenna relays assist in the transmission between a/several multiple-antenna transmitter(s) and a multiple-antenna receiver. Each two nodes are assumed to be either connected through a quasi-static Rayleigh fading channel, or disconnected. We propose a new scheme, which we call *random sequential* (RS), based on the amplify-and-forward relaying. We prove that for general multiple-antenna multiple-relay networks, the proposed scheme achieves the maximum diversity gain. Furthermore, we derive diversity-multiplexing tradeoff (DMT) of the proposed RS scheme for general single-antenna multiple-relay networks. It is shown that for single-antenna two-hop multiple-access multiple-relay ($K > 1$) networks (without direct link between the transmitter(s) and the receiver), the proposed RS scheme achieves the optimum DMT. However, for the case of multiple access single relay setup, we show that the RS scheme reduces to the naive amplify-and-forward relaying and is not optimum in terms of DMT, while the dynamic decode-and-forward scheme is shown to be optimum for this scenario [1].

I. Introduction

A. Motivation

In recent years, relay-assisted transmission has gained significant attention as a powerful technique to enhance the performance of wireless networks, combat the fading effect, extend the coverage, and reduce the amount of interference due to frequency reuse. The main idea is to deploy some extra nodes in the network to facilitate the communication between the end terminals. In this manner, these supplementary nodes act as spatially distributed antennas for the end terminals. More recently, cooperative diversity

Financial supports provided by Nortel, and the corresponding matching funds by the Federal government: Natural Sciences and Engineering Research Council of Canada (NSERC) and Province of Ontario: Ontario Centres of Excellence (OCE) are gratefully acknowledged.

[1] A Part of this paper, Theorem 2, is reported in *Library and Archives Canada Technical Report* [1]. Subsequently, [2] covers Theorems 2 and 3 and [3] covers Theorems 2, 3, 5 and 6. The materials of this paper are reported in [4].

techniques have been proposed as candidates to exploit the spatial diversity offered by the relay networks (for example, see [5]–[8]). A fundamental measure to evaluate the performance of the existing cooperative diversity schemes is the diversity-multiplexing tradeoff (DMT) which was first introduced by Zheng and Tse in the context of point-to-point MIMO fading channels [9]. Roughly speaking, the diversity-multiplexing tradeoff identifies the optimal compromise between the "transmission reliability" and the"data rate" in the high-SNR regime.

In spite of all the interest in relay networks, none of the existing cooperative diversity schemes is proved to achieve the optimum DMT. The problem has been open even for the simple case of half-duplex single-relay single-source single-destination single-antenna setup. Indeed, the only existing DMT achieving scheme for the single-relay channel reported in [7] requires knowledge of CSI (channel state information) for all the channels at the relay node.

B. Related Works

The DMT of relay networks was first studied by Laneman *et al.* in [5] for half-duplex relays. In this work, the authors prove that the DMT of a network with single-antenna nodes, composed of a single source and a single destination assisted with K half-duplex relays, is upper-bounded by[2]

$$d(r) = (K+1)(1-r)^{+}. \tag{1}$$

This result can be established by applying either the multiple-access or the broadcast cut-set bound [10] on the achievable rate of the system. In spite of its simplicity, this is still the tightest upper-bound on the DMT of the relay networks. The authors in [5] also suggest two protocols based on decode-and-forward (DF) and amplify-and-forward (AF) strategies for a single-relay system with single-antenna nodes. In both protocols, the relay listens to the source during the first half of the frame, and transmits during the second half. To improve the spectral efficiency, the authors propose an incremental relaying protocol in which the receiver sends a single bit feedback to the transmitter and to the relay to clarify if it has decoded the transmitter's message or needs help from the relay for this purpose. However, none of the proposed schemes are able to achieve the DMT upper-bound.

The non-orthogonal amplify-and-forward (NAF) scheme, first proposed by Nabar *et al.* in [11], has been further studied by Azarian *at al.* in [6]. In addition to analyzing the DMT of the NAF scheme, reference [6] shows that NAF is the best in the class of AF strategies for single-antenna single-relay systems. The dynamic decode-and-forward (DDF) scheme has been proposed independently in [6], [12], [13] based on the DF strategy. In DDF, the relay node listens to the sender until it can decode the message, and then

[2]Throughout the paper, for any real value a, $a^{+} \equiv \max\{0, a\}$.

re-encodes and forwards it to the receiver in the remaining time. Reference [6] analyzes the DMT of the DDF scheme and shows that it is optimal for low rates in the sense that it achieves (1) for the multiplexing gains satisfying $r \leq 0.5$. However, for higher rates, the relay should listen to the transmitter for most of the time, reducing the spectral efficiency. Hence, the scheme is unable to follow the upper-bound for high multiplexing gains. More importantly, the generalizations of NAF and DDF for multiple-relay systems fall far from the upper-bound, especially for high multiplexing gains.

Yuksel *et al.* in [7] apply compress-and-forward (CF) strategy and show that CF achieves the DMT upper-bound for multiple-antenna half-duplex single-relay systems. However, in their proposed scheme, the relay node needs to know the CSI of all the channels in the network which may not be practical.

Most recently, Yang *et al.* in [14] propose a class of AF relaying scheme called slotted amplify-and-forward (SAF) for the case of half-duplex multiple-relay ($K > 1$) and single source/destination setup. In SAF, the transmission frame is divided into M equal length slots. In each slot, each relay transmits a linear combination of the previous slots. Reference [14] presents an upper-bound on the DMT of SAF and shows that it is impossible to achieve the MISO upper-bound for finite values of M, even with the assumption of full-duplex relaying. However, as M goes to infinity, the upper-bound meets the MISO upper-bound. Motivated by this upper-bound, the authors in [14] propose a half-duplex sequential SAF scheme. In the sequential SAF scheme, following the first slot, in each subsequent slot, one and only one of the relays is permitted to transmit an amplified version of the signal it has received in the previous slot. By doing this, the different parts of the signal are transmitted through different paths by different relays, resulting in some form of spatial diversity. However, [14] could only show that the sequential SAF achieves the MISO upper-bound for the setup of non-interfering relays, i.e. when the consecutive relays (ordered by transmission times) do not cause any interference on one another.

Apart from investigating the optimum diversity-multiplexing tradeoff for relay networks, recently, other aspects of the relay networks has also been studied (for example, see [15]–[27]). [15], [16] develop new coding schemes based on Decode-and-Forward and Compress-and-Forward relaying strategies for relay networks. Avestimehr *et al.* in [19] study the outage capacity of the relay channel for low-SNR regime and show that in this regime, the bursty Amplify-and-Forward relaying protocol achieves the optimum outage. Avestimehr *et al.* in [20] present a linear deterministic model for the wireless relay network and characterize its exact capacity. Applying the capacity-achieving scheme of the corresponding deterministic model, the authors in [20] show that the capacity of wireless single-relay channel and the diamond relay channel can be characterized within 1 bit and 2 bits, respectively, regardless of the values of the channel gains. The scaling law capacity of large wireless networks is addressed in [21]–[27]. Gastpar *et al.* in [23] prove that employing AF relaying achieves the capacity of the Gaussian parallel single-antenna relay

network for asymptotically large number of relays. Bolcskei *et al.* in [24] extend the work of [23] to the parallel multiple-antenna relay network and characterize the capacity of network within $O(1)$, for large number of relays. Oveis Gharan *et al.* in [25] propose a new AF relaying scheme for parallel multiple-antenna fading relay networks. Applying the proposed AF scheme, the authors in [25] characterize the capacity of parallel multiple-antenna relay networks for the scenario where either the number of relays is large or the power of each relay tends to infinity.

Recently, in a parallel and independent work by Kumar *et al* [28][3] the possibility of achieving the optimum DMT is shown in single-antenna half-duplex relay networks with some graph topologies including KPP , KPP(I), KPP(D) graphs for $K \geq 3$. A KPP graph is a directed graph consisted of K vertex-disjoint paths each with the length greater than one, connecting the transmitter to the receiver. KPP(I) is a directed graph consisted of K vertex-disjoint paths each with length greater than one, connecting the transmitter to the receiver, and possible edges between different paths. KPP(D) is a directed graph consisted of K vertex-disjoint paths each with length greater than one, and a direct path connecting the transmitter to the receiver. It is worth mentioning that in all the mentioned graph topologies, the upper-bound of DMT is achieved by a cut-set of the MISO or SIMO form, i.e. all edges crossing the cut are originated from or destined to the same vertex. Also, they show that the maximum diversity can be achieved in a general multiple-antenna multiple relays network.

C. Contributions

In this paper, we propose a new scheme, which we call random sequential (RS), based on the SAF relaying for general multiple-antenna multi-hop networks. The key elements of the proposed scheme are: 1) signal transmission through sequential paths in the network, 2) path timing such that no non-causal interference is caused from the transmitter of the future paths on the receiver of the current path, 3) multiplication by a random unitary matrix at each relay node, and 4) no signal boosting in amplify-and-forward relaying at the relay nodes, i.e. the received signal is amplified by a coefficient with the absolute value of at most 1. Furthermore, each relay node knows the CSI of its corresponding backward channel, and the receiver knows the equivalent end-to-end channel. We prove that this scheme achieves the maximum diversity gain in a general multiple-antenna multiple-relay network (no restriction imposed on the set of interfering node pairs). Furthermore, we derive the DMT of the RS scheme for general single-antenna multiple-relay networks. Specifically, we derive: 1) the exact DMT of the RS scheme under the condition of "non-interfering relaying", and 2) a lower-bound on the DMT of the RS scheme (no conditions imposed). Finally, we prove that for single-antenna multiple-access multiple-relay networks

[3]After the completion of this work, the authors became aware of [28].

(with $K > 1$ relays) when there is no direct link between the transmitters and the receiver and all the relays are connected to the transmitter and to the receiver, the RS scheme achieves the optimum DMT. However, for two-hop multiple-access single-relay networks, we show that the proposed scheme is unable to achieve the optimum DMT, while the DDF scheme is shown to perform optimum in this scenario.

It is worth mentioning that the optimality results in this paper can easily be applied to the case of KPP and KPP(D) graphs introduced in [28]. However, the proof approach we use in this paper is entirely different from that of used in [28]; Our proofs are based on the matrix inequalities while the proofs of [28] are based on information-theoretic inequalities. Furthermore, [28] shows the achievability of the maximum diversity gain in a general multiple-antenna multiple-relay network by considering a multiple-antenna node as multiple single-antenna nodes and using just one antenna at each time, while in our proof we show that the proposed RS scheme in general can achieve the maximum diversity also in the MIMO form and by using all the antennas simultaneously. Finally, the achievability of the linear DMT between the points $(0, d_{\max})$ and $(1, 0)$ in single-antenna layered network and directed acyclic graph network with full-duplex relays is independently shown as a remark of Theorems 1 and 4 in our paper, respectively.

The rest of the paper is organized as follows. In section II, the system model is introduced. In section III, the proposed random sequential scheme (RS) is described. Section IV is dedicated to the DMT analysis of the proposed RS scheme. Section V proves the optimality of the RS scheme in terms of diversity gain in general multiple-antenna multiple-relay networks. Finally, section VI concludes the paper.

D. Notations

Throughout the paper, the superscripts ^{T} and ^{H} stand for matrix operations of transposition and conjugate transposition, respectively. Capital bold letters represent matrices, while lowercase bold letters and regular letters represent vectors and scalars, respectively. $\|\mathbf{v}\|$ denotes the norm of vector $\mathbf{v}$ while $\|\mathbf{A}\|$ represents the Frobenius norm of matrix $\mathbf{A}$. $|\mathbf{A}|$ denotes the determinant of matrix $\mathbf{A}$. $\log(.)$ denotes the base-2 logarithm. The notation $\mathbf{A} \preccurlyeq \mathbf{B}$ is equivalent to $\mathbf{B} - \mathbf{A}$ is a positive semi-definite matrix. Motivated by the definition in [9], we define the notation $f(P) \doteq g(P)$ as $\lim_{P\to\infty} \frac{f(P)}{\log(P)} = \lim_{P\to\infty} \frac{g(P)}{\log(P)}$. Similarly, $f(P) \dot{\leq} g(P)$ and $f(P) \dot{\geq} g(P)$ are equivalent to $\lim_{P\to\infty} \frac{f(P)}{\log(P)} \leq \lim_{P\to\infty} \frac{g(P)}{\log(P)}$ and $\lim_{P\to\infty} \frac{f(P)}{\log(P)} \geq \lim_{P\to\infty} \frac{g(P)}{\log(P)}$, respectively. Finally, we use $A \approx B$ to denote the approximate equality between A and B, such that by substituting A by B the validity of the equations are not compromised.

II. System Model

Our setup consists of K relays assisting the transmitter and the receiver in the half-duplex mode, i.e. at a given time, the relays can either transmit or receive. Each two nodes are assumed either i) to be connected by a quasi-static flat Rayleigh-fading channel, i.e. the channel gains remain constant during a

block of transmission and change independently from block to block; or ii) to be disconnected, i.e. there is no direct link between them. Hence, the undirected graph $G = (V, E)$ is used to show the connected pairs in the network[4]. The node set is denoted by $V = \{0, 1, \ldots, K+1\}$ where the i'th node is equipped with N_i antennas. Nodes 0 and $K+1$ correspond to the transmitter and the receiver nodes, respectively[5]. The received and the transmitted vectors at the k'th node are shown by $\mathbf{y}_k$ and $\mathbf{x}_k$, respectively. Hence, at the receiver side of the a'th node, we have

$$\mathbf{y}_a = \sum_{\{a,b\} \in E} \mathbf{H}_{a,b} \mathbf{x}_b + \mathbf{n}_a, \tag{2}$$

where $\mathbf{H}_{a,b}$ shows the $N_a \times N_b$ Rayleigh-distributed channel matrix between the a'th and the b'th nodes and $\mathbf{n}_a \sim \mathcal{N}\left(\mathbf{0}, \mathbf{I}_{N_a}\right)$ is the additive white Gaussian noise. We assume reciprocal channels between each two nodes. Hence, $\mathbf{H}_{a,b} = \mathbf{H}_{b,a}^T$. However, it can be easily verified that all the statements of the paper are valid under the non-reciprocity assumption. In the scenario of single-antenna networks, the channel between nodes a and b is denoted by $h_{\{a,b\}}$ to emphasize both the SISO and the reciprocally assumptions. As in [6] and [14], each relay is assumed to know the state of its backward channel, and moreover, the receiver knows the equivalent end-to-end channel. Hence, unlike the CF scheme in [7], no CSI feedback is needed. All nodes have the same power constraint, P. Finally, we assume that the topology of the network is known by the nodes such that they can perform a distributed AF strategy throughout the network.

Throughout the section on diversity-multiplexing tradeoff, we make some further assumptions in order to prove our statements. First, we consider the scenario in which nodes with a single antenna are used. Moreover, in Theorems 2, 3, 5, and 6, where we address DMT optimality of the RS scheme, we assume that there is no direct link between the transmitter(s) and the receiver. This assumption is reasonable when the transmitter and the receiver are far from each other and the relay nodes establish the connection between the end nodes. Moreover, we assume that all the relay nodes are connected to the transmitter and to the receiver through quasi-static flat Rayleigh-fading channels. Hence, the network graph is two-hop. In specific, we denote the output vector at the transmitter as $\mathbf{x}$, the input vector and the output vector at the k'th relay as $\mathbf{r}_k$ and $\mathbf{t}_k$, respectively, and the input at the receiver as $\mathbf{y}$.

[4]Note that however, in Remarks 2 and 6, the directed graph is considered.

[5]Throughout the paper, it is assumed that the network consists of one transmitter. However, in Theorems 5 and 6, we study the case of two-hop multiple transmitters single receiver scenario.

III. Proposed Random Sequential (RS) Amplify-and-Forwarding Scheme

In the proposed RS scheme, a sequence $\mathrm{P} \equiv (\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$ of L paths[6] originating from the transmitter and destinating to the receiver with the length $(l_1, l_2, \ldots, l_L)$ are involved in connecting the transmitter to the receiver sequentially ($\mathrm{p}_i(0) = 0, \mathrm{p}_i(l_i) = K+1$). Note that any path p of G can be selected multiple times in the sequence.

Furthermore, the entire block of transmission is divided into S slots, each consisting of T' symbols. Hence, the entire block consists of $T = ST'$ symbols. Let us assume the transmitter intends to send information to the receiver at a rate of r bits per symbol. To transmit a message w, the transmitter selects the corresponding codeword from a Gaussian random code-book consisting of $2^{ST'r}$ elements each of with length LT'. Starting from the first slot, the transmitter sequentially transmits the i'th portion ($1 \leq i \leq L$) of the codeword through the sequence of relay nodes in p_i. More precisely, a timing sequence $\{s_{i,j}\}_{i=1,j=1}^{L,l_i}$ is associated with the path sequence. The transmitter sends the i'th portion of the codeword in the $s_{i,1}$'th slot. Following the transmission of the i'th portion of the codeword by the transmitter, in the $s_{i,j}$'th slot, $1 \leq j \leq l_i$, the node $\mathrm{p}_i(j)$ receives the transmitted signal from the node $\mathrm{p}_i(j-1)$. Assuming $\mathrm{p}_i(j)$ is not the receiver node, i.e. $j < l_i$, it multiplies the received signal in the $s_{i,j}$'th slot by a $N_{\mathrm{p}_i(j)} \times N_{\mathrm{p}_i(j)}$ random, uniformly distributed unitary matrix $\mathbf{U}_{i,j}$ which is known at the receiver side, amplifies the signal by the maximum possible coefficient $\alpha_{i,j}$ considering the output power constraint P and $\alpha_{i,j} \leq 1$, and transmits the amplified signal in the $s_{i,j+1}$'th slot. Furthermore, the timing sequence $\{s_{i,j}\}$ should have the following properties

(1) for all i, j, we have $1 \leq s_{i,j} \leq S$.

(2) for $i < i'$, we have $s_{i,1} < s_{i',1}$ (the ordering assumption on the paths)

(3) for $j < j'$, we have $s_{i,j} < s_{i,j'}$ (the causality assumption)

(4) for all $i < i'$ and $s_{i,j} = s_{i',j'}$, we have $\{\mathrm{p}_i(j), \mathrm{p}_{i'}(j'-1)\} \notin E$ (no noncausal interference assumption). This assumption ensures that the signal of the future paths causes no interference on the output signal of the current path. This assumption can be realized by designing the timing of the paths such that in each time slot, the current running paths are established through disjoint hops.

At the receiver side, having received the signal of all paths, the receiver decodes the transmitted message w based on the signal received in the time slots $\{s_{i,l_i}\}_{i=1}^{L}$. As we observe in the sequel, the fourth assumption on $\{s_{i,j}\}$ converts the equivalent end-to-end channel matrix to lower-triangular in the case of

[6]Throughout the paper, a path p is defined as a sequence of the graph nodes $(v_0, v_1, v_2, \ldots, v_l)$ such that for any i, $\{v_i, v_{i+1}\} \in E$, and for all $i \neq j$, we have $v_i \neq v_j$. The length of the path is defined as the total number of edges on the path, l. Furthermore, $\mathrm{p}(i)$ denotes the i'th node that p visits, i.e. $\mathrm{p}(i) = v_i$.

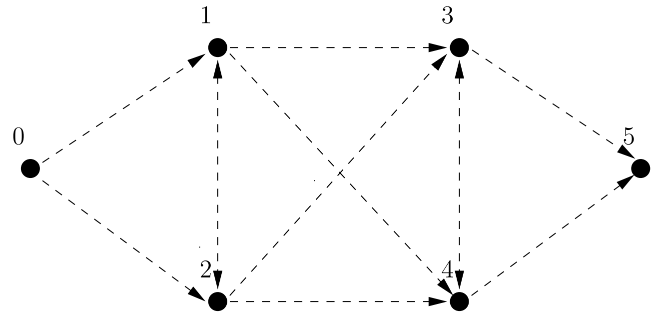

Fig. 1. An example of a 3 hops network where $N_0 = N_5 = 2, N_1 = N_2 = N_3 = N_4 = 1$.

single-antenna nodes, or to block lower-triangular in the case of multiple-antenna nodes.

An example of a three-hop network consisting of $K = 4$ relays is shown in figure (1). It can easily be verified that there are exactly 12 paths in the graph connecting the transmitter to the receiver. Now, consider the four paths $\mathrm{p}_1 = (0, 1, 3, 5)$, $\mathrm{p}_2 = (0, 2, 4, 5)$, $\mathrm{p}_3 = (0, 1, 4, 5)$ and $\mathrm{p}_4 = (0, 2, 3, 5)$ connecting the transmitter to the receiver. Assume the RS scheme is performed with the path sequence $\mathrm{P}_1 \equiv (\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_3, \mathrm{p}_4)$. Table I shows one possible valid timing sequence associated with RS scheme with the path sequence P_1. As seen, the first portion of the transmitter's codeword is sent in the 1st time slot and is received by the receiver through the nodes of the path $\mathrm{P}_1(1) \equiv (0, 1, 3, 5)$ as follows: In the 1st slot, the transmitter's signal is received by node 1. Following that, in the 2nd slot, node 1 sends the amplified signal to node 3, and finally, in the 3rd slot, the receiver receives the signal from node 3. As observed, for every $1 \leq i \leq 3$, signal of the i'th path interferes on the output signal of the $i + 1$'th path. However, no interference is caused by the signal of future paths on the outputs of the current path. The timing sequence corresponding to Table I can be expressed as $s_{i,j} = i + \lfloor \frac{i}{3} \rfloor + j - 1$ and it results in the total number of transmission slots to be equal to 7, i.e. $S = 7$.

As an another example, consider RS scheme with the path sequence $\mathrm{P}_2 \equiv (\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_1, \mathrm{p}_2)$. Table II shows one possible valid timing-sequence for the RS scheme with the path sequence P_2. Here, we observe that the signal on every path interferes on the output of the next two consecutive paths. However, like the

time-slot	1	2	3	4	5	6	7
$\mathrm{P}_1(1)$	$0 \to 1$	$1 \to 3$	$3 \to 5$	—	—	—	—
$\mathrm{P}_1(2)$	—	$0 \to 2$	$2 \to 4$	$4 \to 5$	—	—	—
$\mathrm{P}_1(3)$	—	—	—	$0 \to 1$	$1 \to 4$	$4 \to 5$	—
$\mathrm{P}_1(4)$	—	—	—	—	$0 \to 2$	$2 \to 3$	$3 \to 5$

TABLE I

ONE POSSIBLE VALID TIMING FOR RS SCHEME WITH THE PATH SEQUENCE $\mathrm{P}_1 = (\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_3, \mathrm{p}_4)$.

time-slot	1	2	3	4	5	6
$\mathrm{P}_2(1)$	$0 \to 1$	$1 \to 3$	$3 \to 5$	—	—	—
$\mathrm{P}_2(2)$	—	$0 \to 2$	$2 \to 4$	$4 \to 5$	—	—
$\mathrm{P}_2(3)$	—	—	$0 \to 1$	$1 \to 3$	$3 \to 5$	—
$\mathrm{P}_2(4)$	—	—	—	$0 \to 2$	$2 \to 4$	$4 \to 5$

TABLE II

ONE POSSIBLE VALID TIMING FOR RS SCHEME WITH THE PATH SEQUENCE $\mathrm{P}_2 = (\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_1, \mathrm{p}_2)$.

scenario with P_1, no interference is caused by the signal of future paths on the output signal of the current path. The timing sequence corresponding to Table II can be expressed as $s_{i,j} = i + j - 1$ and it results in the total number of transmission slots equal to 6, i.e. $S = 6$.

It is worth noting that to achieve higher spectral efficiencies (corresponding to larger multiplexing gains), it is desirable to have larger values for $\frac{L}{S}$. Indeed, $\frac{L}{S} \to 1$ is the highest possible value. However, this can not be achieved in some graphs (an example is the case of two-hop single relay scenario studied in the next section where $\frac{L}{S} = 0.5$). On the other hand, to achieve higher reliability (corresponding to larger diversity gains between the end nodes), it is desirable to utilize more paths of the graph in the path sequence. It is not always possible to satisfy both of these objectives simultaneously. As an example, consider the single-antenna two-hop relay network where there is a direct link between the end nodes, i.e. G is the complete graph. Here, all the nodes of the graph interfere on each other, and consequently, in each time slot only one path can transmit signal. Hence, in order to achieve $\frac{L}{S} \to 1$, only the direct path $(0, K+1)$ should be utilized for almost all the time.

As an another example, consider the 3-hop network in figure (1). As we will see in the following sections, the RS scheme corresponding to the path sequence P_1 achieves the maximum diversity gain of the network, $d = 4$. However, it can easily be verified that no valid timing-sequence can achieve fewer number of transmission slots than the one shown in Table I. Hence, $\frac{L}{S} = \frac{4}{7}$ is the best RS scheme can achieve with P_1. On the other hand, consider the RS scheme with the path sequence P_2. Although, as seen in the sequel, the scheme achieves the diversity gain $d = 2$ which is below the maximum diversity gain of the network, it utilizes fewer number of slots compared to the case using the path sequence P_1. Indeed, it achieves $\frac{L}{S} = \frac{4}{6}$.

In the two-hop scenario investigated in the next section, we will see that for asymptotically large values of L, it is possible to utilize all the paths needed to achieve the maximum diversity gain and, at the same time, devise the timing sequence such that $\frac{L}{S} \to 1$. Consequently, it will be shown that in this setup, the proposed RS scheme achieves the optimum DMT.

IV. Diversity-Multiplexing Tradeoff

In this section, we analyze the performance of the RS scheme in terms of the DMT for the single-antenna multiple-relay networks. First, in subsection *A*, we study the performance of the RS scheme for the case of non-interfering relays where there exists neither causal nor noncausal interference between the signals sent through different paths. In this case, as there exists no interference between different paths, we can assume that the amplification coefficients take values greater than one, i.e. the constraint $\alpha_{i,j} \leq 1$ can be omitted. Under the condition of non-interfering relays, we derive the exact DMT of the RS scheme. As a result, we show that the RS scheme achieves the optimum DMT for the setup of non-interfering two-hop multiple-relay ($K > 1$) single-transmitter single-receiver, where there exists no direct link between the relay nodes and between the transmitter and the receiver (more precisely, $E = \{\{0, k\}, \{k, K+1\}\}_{k=1}^{K}$). To prove this, we assume that the RS scheme relies on $L = BK$ paths, $S = BK + 1$ slots, where B is an integer number, and the path sequence is $\mathrm{Q} \equiv (\mathrm{q}_1, \ldots, \mathrm{q}_K, \mathrm{q}_1, \ldots, \mathrm{q}_K, \ldots, \mathrm{q}_1, \ldots, \mathrm{q}_K)$ where $\mathrm{q}_k \equiv (0, k, K+1)$. In other words, every path q_k is used B times in the sequence. Here, each K consecutive slots are called a sub-block. Hence, the entire block of transmission consists of $B+1$ sub-blocks. The timing sequence is defined as $s_{i,j} = i + j - 1$. It is easy to verify that the timing sequence satisfies the requirements. Here, we observe that the spectral efficiency is $\frac{L}{S} = 1 - \frac{1}{S}$ which converges to 1 for asymptotically large values of S. By deriving the exact DMT of the RS scheme, we prove that the RS scheme achieves the optimum DMT for asymptotically large values of S.

In subsection *B*, we study the performance of the RS scheme for general single-antenna multiple-relay networks. First, we study the performance of RS scheme for the setup of two-hop single-transmitter single-receiver multiple-relay ($K > 1$) networks where there exists no direct link between the transmitter and the receiver; However, no additional restriction is imposed on the graph of the interfering relay pairs. We apply the RS scheme with the same parameters used in the case of two-hop non-interfering networks. We derive a lower-bound for DMT of the RS scheme. Interestingly, it turns out that the derived lower-bound merges to the upper-bound on the DMT for asymptotic values of B. Next, we generalize our result and derive a lower-bound on DMT of the RS scheme for general single-antenna multiple-relay networks.

Finally, in subsection *C*, we generalize our results for the scenario of single-antenna two-hop multiple-access multiple-relay ($K > 1$) networks where there exists no direct link between the transmitters and the receiver. Here, we apply the RS scheme with the same parameters as used in the case of single-transmitter single-receiver two-hop relay networks. However, it should be noted that here, instead of sending data from the single transmitter, all the transmitters send data coherently. By deriving a lower-bound on the DMT of the RS scheme, we show that in this network the RS scheme achieves the optimum DMT. However, as studied in subsection *D*, for the setup of single-antenna two-hop multiple-access single-relay

networks where there exists no direct link between the transmitters and the receiver, the proposed RS scheme reduces to naive amplify-and-forward relaying and is not optimum in terms of the DMT. In this setup, we show that the DDF scheme achieves the optimum DMT.

A. Non-Interfering Relays

In this subsection, we study the DMT behavior of the RS scheme in general single-antenna multi-hop relay networks under the condition that there exists neither causal nor noncausal interference between the signals transmitted over different paths. More precisely, we assume the timing sequence is designed such that if $s_{i,j} = s_{i',j'}$, then we have $\{\mathrm{p}_i(j), \mathrm{p}_{i'}(j'-1)\} \notin E$. This assumption is stronger than the fourth assumption on the timing sequence (here the condition $i < i'$ is omitted). We call this the "non-interfering relaying" condition. Under this condition, as there exists no interference between signals over different paths, we can assume that the amplification coefficients take values greater than one, i.e. the constraint $\alpha_{i,j} \leq 1$ can be omitted.

First, we need the following definition.

Definition 1 *For a network with the connectivity graph $G = (V, E)$, a cut-set on G is defined as a subset $\mathcal{S} \subseteq V$ such that $0 \in \mathcal{S}, K+1 \in \mathcal{S}^c$. The weight of the cut-set corresponding to $\mathcal{S}$, denoted by $w(\mathcal{S})$, is defined as*

$$w_G(\mathcal{S}) \quad = \sum_{a \in \mathcal{S}, b \in \mathcal{S}^c, \{a,b\} \in E} N_a \times N_b. \tag{3}$$

Theorem 1 *Consider a half-duplex single-antenna multiple-relay network with the connectivity graph $G = (V, E)$. Assuming "non-interfering relaying", the RS scheme with the path sequence $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$ achieves the diversity gain corresponding to the following linear programming optimization problem*

$$d_{RS,NI}(r) = \min_{\boldsymbol{\mu} \in \hat{\mathcal{R}}} \sum_{e \in E} \mu_e, \tag{4}$$

where $\boldsymbol{\mu}$ is a vector defined on edges of G and $\hat{\mathcal{R}}$ is a region of $\boldsymbol{\mu}$ defined as

$$\hat{\mathcal{R}} \equiv \left\{ \boldsymbol{\mu} \,\middle|\, \mathbf{0} \leq \boldsymbol{\mu} \leq \mathbf{1}, \sum_{i=1}^{L} \max_{1 \leq j \leq l_i} \mu_{\{\mathrm{p}_i(j), \mathrm{p}_i(j-1)\}} \geq L - Sr \right\}.$$

Furthermore, the DMT of the RS scheme can be upper-bounded as

$$d_{RS,NI}(r) \leq (1-r)^+ \min_{\mathcal{S}} w_G(\mathcal{S}), \tag{5}$$

where $\mathcal{S}$ is a cut-set on G. Finally, by properly selecting the path sequence, one can always achieve

$$d_{RS,NI}(r) \geq (1 - l_G r)^+ \min_{\mathcal{S}} w_G(\mathcal{S}), \tag{6}$$

where $\mathcal{S}$ is a cut-set on G and l_G is the maximum path length between the transmitter and the receiver.

Proof: Since the relay nodes are non-interfering, the achievable rate of the RS scheme for a realization of the channels is equal to

$$R_{RS,NI}\left(\{h_e\}_{e\in E}\right) = \frac{1}{S}\sum_{i=1}^{L}\log\left(1+P\prod_{j=1}^{l_i}|\alpha_{i,j}|^2\left|h_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\right|^2\left(1+\sum_{j=1}^{l_i-1}\prod_{k=j}^{l_i-1}|\alpha_{i,k}|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2\right)^{-1}\right), \tag{7}$$

where $\forall j < l_i : \alpha_{i,j} = \sqrt{\frac{P}{1+\left|h_{\{\mathrm{p}_i(j-1),\mathrm{p}_i(j)\}}\right|^2 P}}$ and $\alpha_{i,l_i} = 1$ (since $\mathrm{p}_i(l_i) = K+1$). In deriving the above equation, we have used the fact that as the paths are non-interfering, the achievable rate can be written as the sum of the rates over the paths, noting that the terms $P\prod_{j=1}^{l_i}|\alpha_{i,j}|^2\left|h_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\right|^2$ and $1+\sum_{j=1}^{l_i-1}\prod_{k=j}^{l_i-1}|\alpha_{i,k}|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2$ represent the effective signal power and the noise power over the ith path, respectively. Hence, the probability of outage equals

$$\begin{aligned}
\mathbb{P}\{\mathcal{E}\} &= \mathbb{P}\left\{R_{RS,NI}\left(\{h_e\}_{e\in E}\right)\le r\log(P)\right\}\\
&\overset{(a)}{\doteq} \mathbb{P}\left\{\prod_{i=1}^{L}\max\left\{P^{-1},\min\left\{\left|h_{\{0,\mathrm{p}_i(1)\}}\right|^2\prod_{k=1}^{j}|\alpha_{i,k}|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2\right\}_{j=0}^{l_i-1}\right\}\le P^{Sr-L}\right\}\\
&\overset{(b)}{\doteq} \max_{\substack{t_1,t_2,\ldots,t_L\\1\le t_i\le l_i}}\mathbb{P}\left\{\prod_{i=1}^{L}\max\left\{P^{-1},\left|h_{\{0,\mathrm{p}_i(1)\}}\right|^2\prod_{k=1}^{t_i-1}|\alpha_{i,k}|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2\right\}\le P^{Sr-L}\right\}\\
&\overset{(c)}{\doteq} \max_{\substack{\mathcal{S}_1,\mathcal{S}_2,\ldots,\mathcal{S}_L\\ \mathcal{S}_i\subseteq\{1,2,\ldots,l_i-1\}}}\ \max_{\substack{t_1,t_2,\ldots,t_L\\ \max\{x\in\mathcal{S}_i\}<t_i\le l_i}}\\
&\qquad \mathbb{P}\left\{\prod_{i=1}^{L}\max\left\{P^{-1},P^{|\mathcal{S}_i|}\left|h_{\{\mathrm{p}_i(t_i),\mathrm{p}_i(t_i-1)\}}\right|^2\prod_{k\in\mathcal{S}_i}\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}}\right|^2\right\}\le P^{Sr-L}\right\}.
\end{aligned}\tag{8}$$

Here, (a) follows from the facts that i) $\forall x\ge 0 : \max\{1,x\}\le 1+x\le 2\max\{1,x\}$, which implies that $1+P\Theta\approx\max(1,P\Theta)$, where

$$\Theta \triangleq \prod_{j=1}^{l_i}|\alpha_{i,j}|^2\left|h_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\right|^2\left(1+\sum_{j=1}^{l_i-1}\prod_{k=j}^{l_i-1}|\alpha_{i,k}|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2\right)^{-1},$$

and ii) for all $x_i\ge 0$, $\frac{1}{M}\min\left\{\frac{1}{x_i}\right\}_{i=1}^{M}\le\left(\sum_{i=1}^{M}x_i\right)^{-1}\le\min\left\{\frac{1}{x_i}\right\}_{i=1}^{M}$, which implies that

$$\left(1+\sum_{j=1}^{l_i-1}\prod_{k=j}^{l_i-1}|\alpha_{i,k}|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2\right)^{-1}\approx\min\left(1,\left\{\left(\prod_{k=j}^{l_i-1}|\alpha_{i,k}|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2\right)^{-1}\right\}_{j=1}^{l_i-1}\right).$$

(b) follows from the fact that for any increasing function $f(.)$, we have

$$\max_{1\le i\le M}\mathbb{P}\left\{f(x_i)\le y\right\}\le\mathbb{P}\left\{f\left(\min_{1\le i\le M}x_i\right)\le y\right\}\le M\max_{1\le i\le M}\mathbb{P}\left\{f(x_i)\le y\right\}.$$

(c) follows from the fact that

$$0.5\min\left\{1, P\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}}\right|^2\right\} \leq \left|\alpha_{i,k}h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}}\right|^2 \leq \min\left\{1, P\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}}\right|^2\right\},$$

which implies that $\left|\alpha_{i,k}h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}}\right|^2 \leq \min\left\{1, P\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}}\right|^2\right\}$. In the last line of (8), $\mathcal{S}_i$ denotes the subset of $\{1,2,\cdots,t_i-1\}$ for which $P\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}}\right|^2 \leq 1$.

Assuming $|h_e|^2 = P^{-\mu_e}$, we define the region $\mathcal{R} \subseteq \mathbb{R}^{|E|}$ as the set of points $\boldsymbol{\mu} = [\mu_e]_{e\in E}$ that the outage event occurs. Let us define $\mathcal{R}_+ = \mathcal{R} \cap (\mathbb{R}_+ \cup \{0\})^{|E|}$. As the probability density function diminishes exponentially as $e^{-P^{\mu_e}}$ for positive values of μ_e, we have $\mathbb{P}\{\mathcal{R}_+\} \doteq \mathbb{P}\{\mathcal{R}\}$. Hence, we have

$$\begin{aligned}
\mathbb{P}\{\mathcal{E}\} &\doteq \mathbb{P}\{\mathcal{R}_+\} \\
&\overset{(a)}{\doteq} \max_{\substack{\mathcal{S}_1,\mathcal{S}_2,\ldots,\mathcal{S}_L \\ \mathcal{S}_i\subseteq\{1,2,\ldots,l_i-1\}}} \max_{\substack{t_1,t_2,\ldots,t_L \\ \max\{x\in\mathcal{S}_i\}<t_i\leq l_i}} \mathbb{P}\{\mathcal{R}(\boldsymbol{\mathcal{S}},\mathbf{t})\} \\
&\overset{(b)}{\doteq} \max_{\substack{\mathbf{t} \\ 1\leq t_i\leq l_i}} \mathbb{P}\{\mathcal{R}_0(\mathbf{t})\},
\end{aligned} \tag{9}$$

where

$$\mathcal{R}(\boldsymbol{\mathcal{S}},\mathbf{t}) \equiv \left\{\boldsymbol{\mu}\in(\mathbb{R}_+\cup\{0\})^{|E|} \left| \sum_{i=1}^{L}\min\left\{1, \mu_{\{\mathrm{p}_i(t_i),\mathrm{p}_i(t_i-1)\}} + \sum_{k\in\mathcal{S}_i}\mu_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}} - |\mathcal{S}_i|\right\} \geq L - Sr \right.\right\},$$

$\mathbf{t} = [t_1, t_2, \ldots, t_L]$, $\boldsymbol{\mathcal{S}} = [\mathcal{S}_1,\cdots,\mathcal{S}_L]$, and $\mathcal{R}_0(\mathbf{t}) \equiv \mathcal{R}(\oslash,\oslash,\ldots,\oslash,t_1,t_2,\ldots,t_L)$, in which $\oslash$ denotes the null set. Here, (a) follows from (8). In order to prove (b), we first show that

$$\min\left\{1, \mu_{\{\mathrm{p}_i(t_i),\mathrm{p}_i(t_i-1)\}} + \sum_{k\in\mathcal{S}_i}\mu_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}} - |\mathcal{S}_i|\right\} \leq \max_{t'_i\in\mathcal{S}_i\cup\{t_i\}}\min\left\{1, \mu_{\left\{\mathrm{p}_i(t'_i),\mathrm{p}_i(t'_i-1)\right\}}\right\}. \tag{10}$$

In order to verify (10), consider two possible scenarios: i) for all $t'_i \in \mathcal{S}_i \cup \{t_i\}$, we have $\mu_{\left\{\mathrm{p}_i(t'_i),\mathrm{p}_i(t'_i-1)\right\}} \leq 1$. In this scenario, as in the left hand side of the inequality, we have the summation of $|\mathcal{S}_i|+1$ positive parameters with value less than or equal to 1 subtracted by $|\mathcal{S}_i|$, we conclude that the left hand side of the inequality is less than or equal to $\mu_{\left\{\mathrm{p}_i(t'_i),\mathrm{p}_i(t'_i-1)\right\}}$ for any $t' \in \mathcal{S}_i \cup \{t_i\}$. Hence, (10) is valid; ii) At least for one $t' \in \mathcal{S}_i \cup \{t_i\}$, we have $\mu_{\left\{\mathrm{p}_i(t'_i),\mathrm{p}_i(t'_i-1)\right\}} > 1$. In this scenario, the right hand side of the inequality is equal to 1 and accordingly, (10) is valid. According to (10), we have $\mathcal{R}(\boldsymbol{\mathcal{S}}_1,\mathbf{t}) \subseteq \bigcup_{\substack{\mathbf{t}' \\ t'_i\in\mathcal{S}_i\cup\{t_i\}}} \mathcal{R}_0(\mathbf{t}')$, which results in (b) of (9).

On the other hand, we know that for $\boldsymbol{\mu}^0 \geq \mathbf{0}$, we have $\mathbb{P}\{\boldsymbol{\mu} \geq \boldsymbol{\mu}^0\} \doteq P^{-\mathbf{1}\cdot\boldsymbol{\mu}^0}$. By taking derivative with respect to $\boldsymbol{\mu}$, we have $f_{\boldsymbol{\mu}}(\boldsymbol{\mu}) \doteq P^{-\mathbf{1}\cdot\boldsymbol{\mu}}$. Let us define $l_0 \triangleq \min_{\boldsymbol{\mu}\in\mathcal{R}_0(\mathbf{t})} \mathbf{1}\cdot\boldsymbol{\mu}$ and $\boldsymbol{\mu}_0 \triangleq \arg\min_{\boldsymbol{\mu}\in\mathcal{R}_0(\mathbf{t})} \mathbf{1}\cdot\boldsymbol{\mu}$, $\mathcal{I} \triangleq [0, l_0]^{2K}$, $\mathcal{I}_0^c \triangleq [\mu_0(1),\infty)\times[\mu_0(2),\infty)\times\cdots\times[\mu_0(L),\infty)$ and for $1 \leq i \leq L$, $\mathcal{I}_i^c \triangleq [0,\infty)^{i-1}\times[l_0,\infty)\times[0,\infty)^{L-i}$. It is easy to verify that $\mathcal{I}_0^c \subseteq \mathcal{R}_0(\mathbf{t})$. Hence, we have

$$\begin{aligned}
\mathbb{P}\{\mathcal{R}_0(\mathbf{t})\} &\overset{(a)}{\doteq} \mathbb{P}\{\mathcal{I}_0^c\} + \int_{R_0(\mathbf{t})\bigcap\mathcal{I}} f_{\boldsymbol{\mu}}(\boldsymbol{\mu})\, d\boldsymbol{\mu} + \sum_{i=1}^{L}\mathbb{P}\{\mathcal{R}_0(\mathbf{t})\cap\mathcal{I}_i^c\} \\
&\overset{(b)}{\doteq} P^{-l_0}.
\end{aligned} \tag{11}$$

Here, (a) follows from the facts that i) $\mathbb{P}\left\{\bigcup_{i=1}^{M}\mathcal{A}_i\right\} \doteq \sum_{i=1}^{M}\mathbb{P}\left\{\mathcal{A}_i\right\}$, and ii) $\mathcal{I}_0^c \subseteq \mathcal{R}_0(\mathbf{t})$ and $\mathbb{R}_+^L = \mathcal{I}\bigcup\left(\bigcup_{i=1}^{L}\mathcal{I}_i^c\right)$ which imply that $\mathcal{R}_0(\mathbf{t})$ can be written as $\mathcal{I}_0^c\bigcup(\mathcal{R}_0(\mathbf{t})\bigcap\mathcal{I})\bigcup\left[\bigcup_{i=1}^{M}(\mathcal{R}_0(\mathbf{t})\bigcap\mathcal{I}_i^c)\right]$. (b) follows from the facts that $\mathbb{P}\left\{\mathcal{I}_0^c\right\} = \mathbb{P}\left\{\boldsymbol{\mu} \geq \boldsymbol{\mu}_0\right\} \doteq P^{-l_0}$, $\int_{R_0(\mathbf{t})\bigcap\mathcal{I}} f_{\boldsymbol{\mu}}(\boldsymbol{\mu})\, d\boldsymbol{\mu} \dot{\leq} \text{vol}\left(R_0(\mathbf{t})\bigcap\mathcal{I}\right)P^{-l_0}$, noting that $\text{vol}\left(R_0(\mathbf{t})\bigcap\mathcal{I}\right)$ is a constant number independent of P, and $\mathbb{P}\left\{\mathcal{R}_0(\mathbf{t})\cap\mathcal{I}_i^c\right\} \leq \mathbb{P}\left\{\mathcal{I}_i^c\right\} = P^{-l_0}$. Now, defining $g_{\mathbf{t}}(\boldsymbol{\mu}) = \sum_{i=1}^{L}\min\left\{1, \mu_{\{\mathrm{p}_i(t_i),\mathrm{p}_i(t_i-1)\}}\right\}$ and $\hat{\boldsymbol{\mu}} = [\min\{\mu_e, 1\}]_{e\in E}$, it is easy to verify that $g_{\mathbf{t}}(\hat{\boldsymbol{\mu}}) = g_{\mathbf{t}}(\boldsymbol{\mu})$ and at the same time $\mathbf{1}\cdot\hat{\boldsymbol{\mu}} < \mathbf{1}\cdot\boldsymbol{\mu}$ unless $\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}$. Hence, defining $\hat{g}_{\mathbf{t}}(\boldsymbol{\mu}) = \sum_{i=1}^{L}\mu_{\{\mathrm{p}_i(t_i),\mathrm{p}_i(t_i-1)\}}$, we have

$$d_{RS,NI}(r) = \min_{\substack{\mathbf{t} \\ 1\leq t_i\leq l_i}} \min_{\substack{\boldsymbol{\mu}\geq\mathbf{0} \\ g_{\mathbf{t}}(\boldsymbol{\mu})\geq L-Sr}} \mathbf{1}\cdot\boldsymbol{\mu} = \min_{\substack{\mathbf{t} \\ 1\leq t_i\leq l_i}} \min_{\substack{\mathbf{0}\leq\boldsymbol{\mu}\leq\mathbf{1} \\ \hat{g}_{\mathbf{t}}(\boldsymbol{\mu})\geq L-Sr}} \mathbf{1}\cdot\boldsymbol{\mu} = \min_{\boldsymbol{\mu}\in\hat{\mathcal{R}}} \mathbf{1}\cdot\boldsymbol{\mu}, \tag{12}$$

where $\hat{\mathcal{R}} = \left\{\boldsymbol{\mu}\,\middle|\,\mathbf{0}\leq\boldsymbol{\mu}\leq\mathbf{1}, \sum_{i=1}^{L}\max_{1\leq j\leq l_i}\mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}} \geq L - Sr\right\}$. This proves the first part of the theorem.

Now, let us define $G_{\mathrm{P}} = (V, E_{\mathrm{P}})$ as the subgraph of G consisting of the edges in the path sequence, i.e. $E_{\mathrm{P}} = \left\{\{\mathrm{p}_i(j), \mathrm{p}_i(j-1)\}, \forall i, j : 1 \leq i \leq L, 1 \leq j \leq l_i\right\}$. Assume $\hat{\mathcal{S}} = \underset{\mathcal{S}}{\text{argmin}}\, w_{G_{\mathrm{P}}}(\mathcal{S})$, where $\mathcal{S}$ is a cut-set on G_{P}. We define $\hat{\boldsymbol{\mu}}$ as $\hat{\mu}_e = \frac{(L-Sr)^+}{L}$ for all $e \in E_{\mathrm{P}}$ such that $|e\cap\hat{\mathcal{S}}| = |e\cap\hat{\mathcal{S}}^c| = 1$ and $\hat{\mu}_e = 0$ for the other edges $e \in E$. As all the paths cross the cutset $\hat{\mathcal{S}}$ at least once, it follows that $\max_{1\leq j\leq l_i}\mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}} = \frac{(L-Sr)^+}{L}$, which implies that $\hat{\boldsymbol{\mu}} \in \hat{\mathcal{R}}$. Hence, we have

$$d_{RS,NI}(r) \leq \mathbf{1}\cdot\hat{\boldsymbol{\mu}} = \frac{(L-Sr)^+}{L}\min_{\mathcal{S}} w_{G_{\mathrm{P}}}(\mathcal{S}) \overset{(a)}{\leq} \frac{(L-Sr)^+}{L}\min_{\mathcal{S}} w_G(\mathcal{S}) \overset{(b)}{\leq} (1-r)^+\min_{\mathcal{S}} w_G(\mathcal{S}), \tag{13}$$

where (a) follows from the fact that as G_{P} is a sub-graph of G, we have $\min_{\mathcal{S}} w_{G_{\mathrm{P}}}(\mathcal{S}) \leq \min_{\mathcal{S}} w_G(\mathcal{S})$ and (b) results from $S \geq L$. This proves the second part of the Theorem.

Finally, we prove the lower-bound on the DMT of the RS scheme. Let us define $d_G = \min_{\mathcal{S}} w_G(\mathcal{S})$. Consider the maximum flow algorithm [29] on G from the source node 0 to the sink node $K+1$. According to the Ford-Fulkerson Theorem [29], one can achieve the maximum flow which is equal to the minimum cut of G by the union of elements of a sequence $(\hat{\mathrm{p}}_1, \hat{\mathrm{p}}_2, \ldots, \hat{\mathrm{p}}_{d_G})$ of paths with the lengths $\left(\hat{l}_1, \hat{l}_2, \ldots, \hat{l}_{d_G}\right)$. Now, consider the RS scheme with $L = L_0 d_G$ paths and the path sequence $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$ consisting of the paths that achieve the maximum flow of G such that any path $\hat{\mathrm{p}}_i$ occurs exactly L_0 times in the sequence. Considering $(l_1, l_2, \ldots, l_L)$ as the length sequence, we select the timing sequence as $s_{i,j} = \sum_{k=1}^{i-1} l_k + j$. It is easy to verify that, not only the timing sequence satisfies the 4 requirements needed for the RS scheme, but also the active relays with the timing sequence are non-interfering. Hence, the assumptions of the first part of the theorem are valid. Moreover, we have $S \leq l_G L$. According to (4), the diversity gain of the RS scheme equals

$$d_{RS,NI}(r) = \min_{\boldsymbol{\mu}\in\hat{\mathcal{R}}}\sum_{e\in E}\mu_e. \tag{14}$$

As $\boldsymbol{\mu} \in \hat{\mathcal{R}}$, we have

$$(L - Sr)^{+} \leq \sum_{i=1}^{L} \max_{1 \leq j \leq l_i} \mu_{\{\mathrm{p}_i(j), \mathrm{p}_i(j-1)\}} \overset{(a)}{\leq} L_0 \sum_{e \in E} \mu_e, \tag{15}$$

where (a) results from the fact that as $(\hat{\mathrm{p}}_1, \hat{\mathrm{p}}_2, \dots, \hat{\mathrm{p}}_{d_G})$ form a valid flow on G (they are non-intersecting over E), every $e \in E$ occurs in at most one $\hat{\mathrm{p}}_i$, or equivalently, in at most L_0 number of p_i's. Combining (14) and (15), we have

$$d_{RS,NI}(r) \geq \frac{(L - Sr)^{+}}{L_0} \geq (1 - l_G r)^{+} d_G = (1 - l_G r)^{+} \min_{\mathcal{S}} w_G(\mathcal{S}). \tag{16}$$

This proves the third part of the Theorem. ■

Remark 1- In scenarios where the minimum-cut on G is achieved by a cut of the MISO or SIMO form, i.e., the edges that cross the cut are either originated from or destined to the same vertex, the upper-bound on the diversity gain of the RS scheme derived in (5) meets the information-theoretic upper-bound on the diversity gain of the network. Hence, in this scenario, any RS scheme that achieves (5) indeed achieves the optimum DMT.

Remark 2- In general, the upper-bound (5) can be achieved for various certain graph topologies by wisely designing the path sequence and the timing sequence. One example is the case of the layered network [20] in which all the paths from the source to the destination have the same length l_G. Let us assume that the relays are allowed to operate in the full-duplex manner. In this case, it easily can be observed that the timing sequence corresponding to the path sequence $(\mathrm{p}_1, \mathrm{p}_2, \dots, \mathrm{p}_L)$ used in the proof of (6) can be modified to $s_{i,j} = i + j - 1$. Accordingly, the number of slots is decreased to $S = L + l_G - 1$. Rewriting (16), we have $d_{RS,NI}(r) = \left(1 - r - \frac{l_G - 1}{L} r\right)^{+} \min_{\mathcal{S}} w_G(\mathcal{S})$ which achieves $(1 - r)^{+} \min_{\mathcal{S}} w_G(\mathcal{S})$ for large values of L.

Next, using Theorem 1, we show that the RS scheme achieves the optimum DMT in the setup of single-antenna two-hop multiple-relay networks where there exists no direct link neither between the transmitter and the receiver, nor between the relay nodes.

Theorem 2 *Assume a single-antenna half-duplex parallel relay scenario with K non-interfering relays. The proposed SM scheme with $L = BK$, $S = BK + 1$, the path sequence*

$$\mathrm{Q} \equiv (\mathrm{q}_1, \dots, \mathrm{q}_K, \mathrm{q}_1, \dots, \mathrm{q}_K, \dots, \mathrm{q}_1, \dots, \mathrm{q}_K)$$

where $\mathrm{q}_k \equiv (0, k, K+1)$ and the timing sequence $s_{i,j} = i + j - 1$ achieves the diversity gain

$$d_{RS,NI}(r) = \max\left\{0, K(1 - r) - \frac{r}{B}\right\}, \tag{17}$$

which achieves the optimum DMT curve $d_{opt}(r) = K(1 - r)^{+}$ as $B \to \infty$.

Proof: First, according to the cut-set bound theorem [10], the point-to-point capacity of the uplink channel (the channel from the transmitter to the relays) is an upper-bound on the achievable rate of the network. Accordingly, the diversity-multiplexing curve of a $1 \times K$ SIMO system which is a straight line (from the multiplexing gain 1 to the diversity gain K, i.e. $d_{opt}(r) = K(1-r)^+$) is an upper-bound on the DMT of the network. Now, we prove that the proposed RS scheme achieves the upper-bound on the DMT for asymptotically large values of S.

As the relay pairs are non-interfering ($1 \leq k \leq K : \{k, (k \mod K) + 1\} \notin E$), the result of Theorem 1 can be applied. As a result

$$d_{RS,NI}(r) = \min_{\boldsymbol{\mu} \in \hat{\mathcal{R}}} \sum_{e \in E} \mu_e, \tag{18}$$

where $\hat{\mathcal{R}} = \left\{ \boldsymbol{\mu} \,\middle|\, \mathbf{0} \leq \boldsymbol{\mu} \leq \mathbf{1}, \sum_{i=1}^{BK} \max_{1 \leq j \leq 2} \mu_{\left\{\mathrm{q}_{(i-1) \mod K+1}(j), \mathrm{q}_{(i-1) \mod K+1}(j-1)\right\}} \geq BK - (BK+1)r \right\}$. Hence, we have

$$BK\left(1 - r - \frac{1}{BK}r\right)^+ \overset{(a)}{\leq} B\sum_{k=1}^{K} \max\left\{\mu_{\{0,k\}}, \mu_{\{K+1,k\}}\right\} \leq B \sum_{e \in E} \mu_e, \tag{19}$$

where (a) results from the fact that every path q_k is used B times in the path sequence. Hence, DMT can be lower-bounded as

$$d_{RS,NI}(r) \geq K\left(1 - r - \frac{1}{BK}r\right)^+. \tag{20}$$

On the other hand, considering the vector $\hat{\boldsymbol{\mu}} = [\hat{\mu}_e]_{e \in E}$ where $\forall 1 \leq k \leq K : \hat{\mu}_{\{0,k\}} = \left(1 - r - \frac{1}{BK}r\right)^+$ and $\forall k, k' \neq 0 : \hat{\mu}_{\{k,k'\}} = 0$, it is easy to verify that $\hat{\boldsymbol{\mu}} \in \hat{\mathcal{R}}$. Hence,

$$d_{RS,NI}(r) \leq \sum_{e \in E} \hat{\mu}_e = K\left(1 - r - \frac{1}{BK}r\right)^+. \tag{21}$$

Combining (20) and (21) completes the proof. ■

Remark 3- Note that as long as the complement[7] of the induced sub-graph of G on the relay nodes $\{1, 2, \ldots, K\}$ includes a Hamiltonian cycle [8], the result of Theorem 2 remains valid. However, the paths $\mathrm{q}_1, \mathrm{q}_2, \ldots, \mathrm{q}_K$ should be permuted in the path sequence according to their orderings in the corresponding Hamiltonian cycle.

According to (17), we observe that the RS scheme achieves the maximum multiplexing gain $1 - \frac{1}{BK+1}$ and the maximum diversity gain K, respectively, for the setup of non-interfering relays. Hence, it achieves the maximum diversity gain for any finite value of B. Also, knowing that no signal is sent to the receiver in the first slot, the RS scheme achieves the maximum possible multiplexing gain. Figure (2) shows the DMT of the scheme for the case of non-interfering relays and various values of K and B.

[7]For every undirected graph $G = (V, E)$, the complement of G is a graph H on the same vertices such that two vertices of H are adjacent if and only if they are non-adjacent in G. [29]

[8]A Hamiltonian cycle is a simple cycle $(v_1, v_2, \cdots, v_K, v_1)$ that goes exactly one time through each vertex of the graph [29].

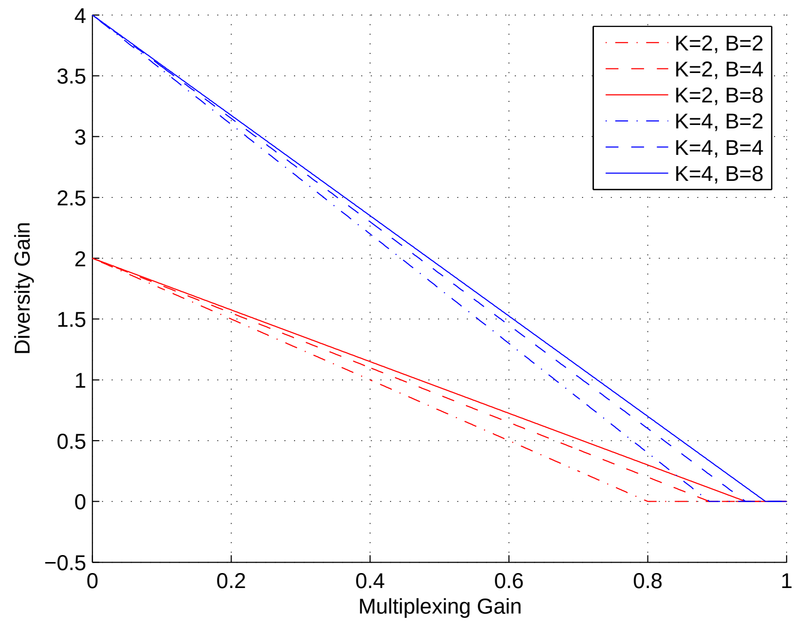

Fig. 2. DMT of RS scheme in parallel relay network for both "interfering" and "non-interfering" relaying scenarios and for different values of K, B.

B. General Case

In this section, we study the performance of the RS scheme in general single-antenna multi-hop wireless networks and derive a lower bound on the corresponding DMT. First, we show that the RS scheme with the parameters defined in Theorem 2 achieves the optimum DMT for the single-antenna parallel-relay networks when there is no direct link between the transmitter and the receiver. Then, we generalize the statement and provide a lower-bound on the DMT of the RS scheme for the more general case.

As stated in the section "System Model", throughout the two-hop network analysis, we slightly modify our notations to simplify the derivations. Specifically, the output vector at the transmitter, the input and the output vectors at the k'th relay, and the input vector at the receiver are denoted as $\mathbf{x}$, $\mathbf{r}_k$, $\mathbf{t}_k$ and $\mathbf{y}$, respectively. h_k and g_k represent the channel gain between the transmitter and the k'th relay and the channel gain between the k'th relay and the destination, respectively. (k) and (b) are defined as $(k) \equiv ((k-2) \mod K) + 1$ and $(b) \equiv b - \lfloor \frac{(k)}{K} \rfloor$. Finally, $i_{(k)}$, $\mathbf{n}_k$, $\mathbf{z}$, and α_k denote the channel gain between the k'th and the (k)'th relay nodes, the noise at the k'th relay and at the receiver, and the amplification coefficient at the k'th relay.

Figure (3) shows a realization of this setup with 4 relays. As observed, the relay set $\{1, 2\}$ is disconnected from the relay set $\{3, 4\}$. In general, the output signal of any relay node k' such that $\{k, k'\} \in E$ can interfere on the received signal of relay node k. However, in Theorem 3, the RS scheme is applied with the same parameters as in Theorem 2. Hence, when the transmitter is sending signal to the k'th relay in

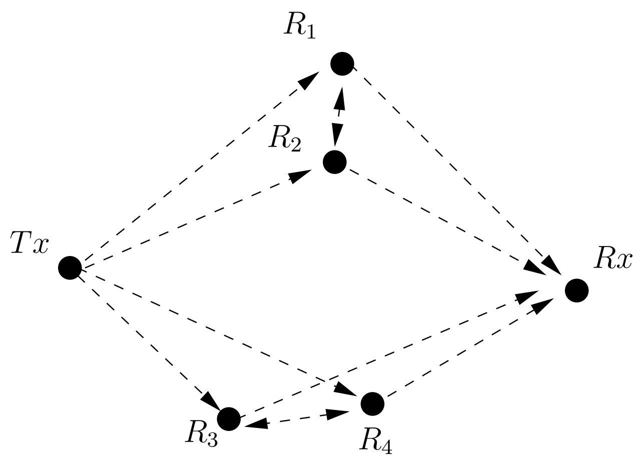

Fig. 3. An example of the half-duplex parallel relay network setup, relay nodes $\{1, 2\}$ are disconnected from relay nodes $\{3, 4\}$.

a time-slot, just the (k)'th relay is simultaneously transmitting and interferes at the k'th relay side. As an example, for the scenario shown in figure (3), we have

$$\begin{aligned}\mathbf{r}_1 &= h_1\mathbf{x} + i_4\mathbf{t}_4 + \mathbf{n}_1, \\ \mathbf{r}_2 &= h_2\mathbf{x} + \mathbf{n}_2.\end{aligned}$$

However, for the sake of simplicity, in the proof of the following theorem, we assume that all the relays interfere with each other. Hence, at the k'th relay, we have

$$\mathbf{r}_k = h_k\mathbf{x} + i_{(k)}\mathbf{t}_{(k)} + \mathbf{n}_k. \tag{22}$$

According to the output power constraint, the amplification coefficient is bounded as $\alpha_k \leq \sqrt{\frac{P}{P\left(|h_k|^2 + \left|i_{(k)}\right|^2\right)+1}}$. However, according to the signal boosting constraint imposed on the RS scheme, we have $|\alpha_k| \leq 1$. Hence, the amplification coefficient is equal to

$$\alpha_k = \min\left\{1, \sqrt{\frac{P}{P\left(|h_k|^2 + \left|i_{(k)}\right|^2\right) + 1}}\right\}. \tag{23}$$

In this manner, it is guaranteed that the noise terms of the different relays are not boosted throughout the network. This is achieved at the cost of working with the output power less than P. On the other hand, we know that almost surely [9] $|h_k|^2, \left|i_{(k)}\right|^2 \dot{\leq} 1$. Hence, almost surely, we have $\alpha_k \doteq 1$. This point will be elaborated further in the proof of the theorem. Now, we prove the DMT optimality of the RS scheme for general single-antenna parallel-relay networks.

[9]By almost surely, we mean its probability is greater than $1 - P^{-\delta}$, for any value of $\delta > 0$.

Theorem 3 *Consider a single-antenna half-duplex parallel relay network with $K > 1$ interfering relays where there is no direct link between the transmitter and the receiver. The diversity gain of the RS scheme with the parameters defined in Theorem 2 is lower-bounded as*

$$d_{RS,I}(r) \geq \max\left\{0, K\left(1-r\right) - \frac{r}{B}\right\}. \tag{24}$$

Furthermore, the RS scheme achieves the optimum DMT $d_{opt}(r) = K(1-r)^+$ as $B \to \infty$.

Proof: First, we show that the entire channel matrix is equivalent to a lower triangular matrix. Let us define $\mathbf{x}_{b,k}, \mathbf{n}_{b,k}, \mathbf{r}_{b,k}, \mathbf{t}_{b,k}, \mathbf{z}_{b,k}, \mathbf{y}_{b,k}$ as the portion of signals that is sent or received in the k'th slot of the b'th sub-block. At the receiver side, we have

$$\begin{aligned} \mathbf{y}_{b,k} &= g_{(k)}\mathbf{t}_{b,k} + \mathbf{z}_{b,k} \\ &= g_{(k)}\alpha_{(k)}\left(\sum_{\substack{1\leq b_1\leq b, 1\leq k_1\leq K \\ b_1K+k_1<bK+k}} p_{b-b_1,k,k_1}\left(h_{k_1}\mathbf{x}_{b_1,k_1} + \mathbf{n}_{b_1,k_1}\right)\right) + \mathbf{z}_{b,k}. \end{aligned} \tag{25}$$

Here, p_{b,k,k_1} has the following recursive formula $p_{0,k,k} = 1, p_{b,k,k_1} = i_{((k))}\alpha_{((k))}p_{(b),(k),k_1}$. Defining the square $BK \times BK$ matrices $\mathbf{G} = \mathbf{I}_B \otimes \textit{diag}\left\{g_1, g_2, \cdots, g_K\right\}$, $\mathbf{H} = \mathbf{I}_B \otimes \textit{diag}\left\{h_1, h_2, \cdots, h_K\right\}$, $\mathbf{\Omega} = \mathbf{I}_B \otimes \textit{diag}\left\{\alpha_1, \alpha_2, \cdots, \alpha_K\right\}$, and

$$\mathbf{F} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ p_{0,2,1} & 1 & 0 & 0 & \cdots \\ p_{0,3,1} & p_{0,3,2} & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ p_{B-1,K,1} & p_{B-1,K,2} & \cdots & p_{0,K,K-1} & 1 \end{pmatrix}, \tag{26}$$

where $\otimes$ is the Kronecker product [30] of matrices and $\mathbf{I}_B$ is the $B \times B$ identity matrix, and the $BK \times 1$ vectors $\mathbf{x}\left(s\right) = \left[x_{1,1}(s), x_{1,2}(s), \cdots, x_{B,K}(s)\right]^T$, $\mathbf{n}\left(s\right) = \left[n_{1,1}\left(s\right), n_{1,2}(s), \cdots, n_{B,K}(s)\right]^T$, $\mathbf{z}\left(s\right) = \left[z_{1,2}(s), z_{1,3}(s), \cdots, z_{B+1,1}(s)\right]^T$, and $\mathbf{y}\left(s\right) = \left[y_{1,2}(s), y_{1,3}(s), \cdots, y_{B+1,1}(s)\right]^T$, we have

$$\mathbf{y}\left(s\right) = \mathbf{G\Omega F}\left(\mathbf{Hx}\left(s\right) + \mathbf{n}\left(s\right)\right) + \mathbf{z}\left(s\right). \tag{27}$$

Here, we observe that the matrix of the entire channel is equivalent to a lower triangular matrix of size $BK \times BK$ for a MIMO system with a colored noise. The probability of outage of such a channel for the multiplexing gain r ($r \leq 1$) is defined as

$$\mathbb{P}\left\{\mathcal{E}\right\} = \mathbb{P}\left\{\log\left|\mathbf{I}_{BK} + P\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right| \leq (BK+1)r\log\left(P\right)\right\}, \tag{28}$$

where $\mathbf{P}_n = \mathbf{I}_{BK} + \mathbf{G\Omega FF}^H\mathbf{\Omega}^H\mathbf{G}^H$, and $\mathbf{H}_T = \mathbf{G\Omega FH}$. Assume $|h_k|^2 = P^{-\mu_k}$, $|g_k|^2 = P^{-\nu_k}$, $|i_k|^2 = P^{-\omega_k}$, and $\mathcal{R}$ as the region in $\mathbb{R}^{3K}$ that defines the outage event $\mathcal{E}$ in terms of the vector $[\boldsymbol{\mu}^T, \boldsymbol{\nu}^T, \boldsymbol{\omega}^T]^T$, where $\boldsymbol{\mu} = \left[\mu_1\mu_2\cdots\mu_K\right]^T, \boldsymbol{\nu} = \left[\nu_1\nu_2\cdots\nu_K\right]^T, \boldsymbol{\omega} = \left[\omega_1\omega_2\cdots\omega_K\right]^T$. The probability distribution function

(and also the complement of the cumulative distribution function) decays exponentially as $P^{-P^{-\delta}}$ for positive values of δ. Hence, the outage region $\mathcal{R}$ is almost surely equal to $\mathcal{R}_+ = \mathcal{R} \bigcap \mathbb{R}_+^{3K}$. Now, we have

$$\begin{aligned}
\mathbb{P}\left\{\mathcal{E}\right\} &\overset{(a)}{\leq} \mathbb{P}\left\{\left|\mathbf{H}_T\right|^2\left|\mathbf{P}_n\right|^{-1} \leq P^{-BK(1-r)+r}\right\} \\
&\overset{(b)}{\leq} \mathbb{P}\left\{-B\sum_{k=1}^{K}\left(\mu_k+\nu_k-\min\left\{0,\mu_k,\omega_{(k)}\right\}\right) - \frac{BK\log(3)+\log\left|\mathbf{P}_n\right|}{\log\left(P\right)} \leq -BK(1-r)+r\right\} \\
&\overset{(c)}{\dot{\leq}} \mathbb{P}\left\{-BK\frac{\log\left[3\left(B^2K^2+1\right)\right]}{\log(P)} + BK\left(1-r\right)-r \leq B\sum_{k=1}^{K}\left(\mu_k+\nu_k\right), \mu_k,\nu_k,\omega_k \geq 0\right\}. \quad (29)
\end{aligned}$$

Here, (a) follows from the fact that for a positive semidefinite matrix $\mathbf{A}$, we have $|\mathbf{I}+\mathbf{A}| \geq |\mathbf{A}|$ and (b) follows from the fact that

$$|\alpha_k|^2 = \min\left\{1, \frac{P}{P^{1-\mu_k}+P^{1-\omega_{(k)}}+1}\right\} \geq \frac{1}{3}\min\left\{1,P,P^{\mu_k},P^{\omega_{(k)}}\right\}$$

and assuming P is large enough such that $P \geq 1$. Finally, (c) is proved as follows:

As $|\alpha_k| \leq 1$, we conclude $p_{n,k,k_1} \leq 1$. Hence, the sum of the entries of each row in $\mathbf{F}\mathbf{F}^H$ is less than B^2K^2. Now, consider the matrix $\mathbf{A} \triangleq B^2K^2\mathbf{I} - \mathbf{F}\mathbf{F}^H$. From the above discussion, it follows that for every i, we have $A_{i,i} \geq \sum_{i\neq j}|A_{i,j}|$. Hence, for every vector $\mathbf{x}$, we have $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq \sum_{i<j}|A_{i,j}|x_i^2 + |A_{i,j}|x_j^2 \pm 2|A_{i,j}|x_ix_j = \sum_{i<j}|A_{i,j}|\left(x_i \pm x_j\right)^2 \geq 0$, and as a result $\mathbf{A}$ is positive semidefinite, which implies that $\mathbf{F}\mathbf{F}^H \preccurlyeq B^2K^2\mathbf{I}_{BK}$. Consequently, we have $\mathbf{P}_n \preccurlyeq \mathbf{I}_{BK} + B^2K^2\mathbf{G}\mathbf{\Omega}\mathbf{\Omega}^H\mathbf{G}^H$. Moreover, Knowing the fact that $\mathbb{P}\left\{\mathcal{R}\right\} \doteq \mathbb{P}\left\{\mathcal{R}_+\right\}$, and conditioned on $\mathcal{R}_+$, one has $|g_k|^2 \leq 1$, which implies that $\mathbf{G}\mathbf{G}^H \preccurlyeq \mathbf{I}$. Combining this with the fact that $\mathbf{\Omega}\mathbf{\Omega}^H \preccurlyeq \mathbf{I}$ (as $|\alpha_k|^2 \leq 1$, $\forall k$) yields $\mathbf{P}_n \preccurlyeq \mathbf{I}_{BK} + B^2K^2\mathbf{G}\mathbf{\Omega}\mathbf{\Omega}^H\mathbf{G}^H \preccurlyeq \left(B^2K^2+1\right)\mathbf{I}_{BK}$. Moreover, conditioned on $\mathcal{R}_+$, we have $\min\left\{0,\mu_k,\omega_{(k)}\right\} = 0$. This completes the proof of (c).

On the other hand, for vectors $\boldsymbol{\mu}^0, \boldsymbol{\nu}^0, \boldsymbol{\omega}^0 \geq \mathbf{0}$, we have $\mathbb{P}\left\{\boldsymbol{\mu} \geq \boldsymbol{\mu}^0, \boldsymbol{\nu} \geq \boldsymbol{\nu}^0, \boldsymbol{\omega} \geq \boldsymbol{\omega}^0\right\} \doteq P^{-\mathbf{1}\cdot\left(\boldsymbol{\mu}^0+\boldsymbol{\nu}^0+\boldsymbol{\omega}^0\right)}$. Similar to the proof of Theorem 1, by taking derivative with respect to $\boldsymbol{\mu},\boldsymbol{\nu}$, we have $f_{\boldsymbol{\mu},\boldsymbol{\nu}}(\boldsymbol{\mu},\boldsymbol{\nu}) \doteq P^{-\mathbf{1}\cdot(\boldsymbol{\mu}+\boldsymbol{\nu})}$. Defining $l_0 \triangleq -\frac{\log\left[3\left(B^2K^2+1\right)\right]}{\log(P)} + (1-r) - \frac{r}{BK}$, $\hat{\mathcal{R}} \triangleq \left\{\boldsymbol{\mu},\boldsymbol{\nu} \geq \mathbf{0}, \frac{1}{K}\mathbf{1}\cdot(\boldsymbol{\mu}+\boldsymbol{\nu}) \geq l_0\right\}$, the cube $\mathcal{I}$ as $\mathcal{I} \triangleq [0,Kl_0]^{2K}$, and for $1 \leq i \leq 2K$, $\mathcal{I}_i^c \triangleq [0,\infty)^{i-1} \times [Kl_0,\infty) \times [0,\infty)^{2K-i}$, we observe

$$\begin{aligned}
\mathbb{P}\left\{\mathcal{E}\right\} &\overset{(a)}{\dot{\leq}} \mathbb{P}\{\hat{\mathcal{R}}\} \\
&\overset{(b)}{\leq} \int_{\hat{\mathcal{R}}\bigcap\mathcal{I}} f_{\boldsymbol{\mu},\boldsymbol{\nu}}\left(\boldsymbol{\mu},\boldsymbol{\nu}\right)d\boldsymbol{\mu}d\boldsymbol{\nu} + \sum_{i=1}^{2K}\mathbb{P}\left\{[\boldsymbol{\mu}^T,\boldsymbol{\nu}^T]^T \in \hat{\mathcal{R}} \cap \mathcal{I}_i^c\right\} \\
&\dot{\leq} \text{vol}(\hat{\mathcal{R}}\cap\mathcal{I})P^{-\min_{\left[\boldsymbol{\mu}_0^T,\boldsymbol{\nu}_0^T\right]^T \in \hat{\mathcal{R}}\bigcap\mathcal{I}} \mathbf{1}\cdot\left(\boldsymbol{\mu}_0+\boldsymbol{\nu}_0\right)} + 2KP^{-Kl_0} \\
&\overset{(c)}{\doteq} P^{-Kl_0} \\
&\doteq P^{-\left[K(1-r)-\frac{r}{B}\right]}. \qquad (30)
\end{aligned}$$

Here, (a) follows from (29), (b) results from writing $\hat{\mathcal{R}}$ as $\left(\hat{\mathcal{R}}\bigcap\mathcal{I}\right)\bigcup\left[\bigcup_{i=1}^{M}\left(\hat{\mathcal{R}}\bigcap\mathcal{I}_i^c\right)\right]$ and using the union bound on the probability, and (c) follows from the fact that $\hat{\mathcal{R}}\bigcap\mathcal{I}$ is a bounded region whose volume is independent of P. (30) completes the proof of Theorem 3. ■

Remark 4- The argument in Theorem 3 is valid no matter what the induced graph of G on the relay nodes is. More precisely, the DMT of the RS scheme can be lower-bounded as (24) as long as $\{0, K+1\} \notin E$ and $\{0,k\}, \{K+1,k\} \in E$. One special case is that the complement of the induced subgraph of G on the relay nodes includes a Hamiltonian cycle which is analyzed in Theorem 2. Here, we observe that the lower-bound on DMT derived in (24) is tight as shown in Theorem 2.

Figure (2) shows the DMT of the RS scheme for varying number of K and B. Noting the proof of Theorem 3, we can easily generalize the result of Theorem 3 and provide a lower-bound on the DMT of the RS scheme for general single-antenna multi-hop multiple-relay networks.

Theorem 4 *Consider a half-duplex single-antenna multiple-relay network with the connectivity graph* $G=(V,E)$ *operated under the RS scheme with* L *paths,* S *slots, and the path sequence* $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$. *Defining* β_e *for each* $e \in E$ *as the number of paths in the path sequence that go through* e*, then the DMT of the RS scheme is lower-bounded as*

$$d_{RS}(r) \geq \frac{L}{\max\limits_{e\in E}\beta_e}\left(1-\frac{S}{L}r\right)^{+}. \tag{31}$$

Proof: First, similar to the proof of Theorem 3, we show that the entire channel matrix is lower triangular. At the receiver side, we have

$$\mathbf{y}_{K+1,i} = \prod_{j=1}^{l_i} h_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\alpha_{i,j}\mathbf{x}_{0,i} + \sum_{j<i} f_{i,j}\mathbf{x}_{0,j} + \sum_{j\leq i, m\leq l_j} q_{i,j,m}\mathbf{n}_{j,m}. \tag{32}$$

Here, $\mathbf{x}_{0,i}$ is the vector transmitted at the transmitter side during the $s_{i,1}$'th slot as the input for the i'th path, $\mathbf{y}_{K+1,i}$ is the vector received at the receiver side during the s_{i,l_i}'th slot as the output for i'th path, $f_{i,j}$ is the interference coefficient which relates the input of the j'th path ($j<i$) to the output of the i'th path, $\mathbf{n}_{j,m}$ is the noise vector during the $s_{j,m}$'th slot at the $\mathrm{p}_j(m)$'th node, and finally, $q_{i,k,m}$ is the coefficient which relates $\mathbf{n}_{k,m}$ to $\mathbf{y}_{K+1,i}$. Note that as the timing sequence satisfies the noncausal interference assumption, the summation terms in (32) do not exceed i. Moreover, for the sake of brevity, we define $\alpha_{i,l_i}=1$. Defining $\mathbf{x}(s) = [x_{0,1}(s)\, x_{0,2}(s) \cdots x_{0,L}(s)]^T$, $\mathbf{y}(s) = [y_{K+1,1}(s)\, y_{K+1,2}(s) \cdots y_{K+1,L}(s)]^T$, and $\mathbf{n}(s) = [n_{1,1}(s)\, n_{1,2}(s) \cdots n_{L,l_L}(s)]^T$, we have the following equivalent lower-triangular matrix between the end nodes:

$$\mathbf{y}(s) = \mathbf{H}_T\mathbf{x}(s) + \mathbf{Q}\mathbf{n}(s). \tag{33}$$

Here,

$$\mathbf{H}_T = \begin{pmatrix} f_{1,1} & 0 & 0 & \dots \\ f_{2,1} & f_{2,2} & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \\ f_{L,1} & f_{L,2} & \dots & f_{L,L} \end{pmatrix}, \tag{34}$$

where $f_{i,i} = \prod_{j=1}^{l_i} h_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\alpha_{i,j}$, and

$$\mathbf{Q} = \begin{pmatrix} q_{1,1,1} & \dots & q_{1,1,l_1} & 0 & 0 & 0 & \dots \\ q_{2,1,1} & \dots & q_{2,1,l_1} & \dots & q_{2,2,l_2} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ q_{L,1,1} & q_{L,1,2} & \dots & \dots & \dots & q_{L,L,l_L-1} & q_{L,L,l_L} \end{pmatrix}. \tag{35}$$

Let us define μ_e for every $e \in E$ such that $|h_e|^2 = P^{-\mu_e}$. First, we observe that similar to the proof of Theorem 3, it can be shown that i) $\alpha_{i,j} \doteq 1$ with probability 1[10], ii) we can restrict ourselves to the region $\mathbb{R}_+$, i.e., the region $\boldsymbol{\mu} > \mathbf{0}$. These two facts imply that $|q_{i,j,m}| \dot{\leq} 1$. This means there exists a constant c which depends just on the topology of the graph G and the path sequence such that $\mathbf{P}_n \triangleq \mathbf{Q}\mathbf{Q}^H \preccurlyeq c\mathbf{I}_L$ (by a similar argument as in the proof of Theorem 3). Hence, similar to the arguments in the equation series (29), the outage probability can be bounded as

$$\begin{aligned} \mathbb{P}\{\mathcal{E}\} &= \mathbb{P}\left\{\left|\mathbf{I}_L + P\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right| \leq P^{Sr}\right\} \\ &\dot{\leq} \mathbb{P}\left\{|\mathbf{H}_T|\left|\mathbf{H}_T^H\right| \leq P^{Sr-L}\right\} \\ &= \mathbb{P}\left\{\sum_{e\in E}\beta_e\mu_e \geq L - Sr\right\} \\ &\doteq \mathbb{P}\left\{\boldsymbol{\mu} \geq \mathbf{0}, \sum_{e\in E}\beta_e\mu_e \geq (L-Sr)^+\right\}, \end{aligned} \tag{36}$$

where β_e is the number of paths in the path sequence that pass through e. Knowing that $\mathbb{P}\{\boldsymbol{\mu} \geq \boldsymbol{\mu}^0\} \doteq P^{-\mathbf{1}\cdot\boldsymbol{\mu}}$ and computing the derivative, we have $f_{\boldsymbol{\mu}}(\boldsymbol{\mu}) = P^{-\mathbf{1}\cdot\boldsymbol{\mu}}$. Defining $\mathcal{R} = \left\{\boldsymbol{\mu} > \mathbf{0}, \sum_{e\in E}\beta_e\mu_e \geq (L-Sr)^+\right\}$ and applying the results of equation series (30), we obtain

$$\mathbb{P}\{\mathcal{E}\} \;\dot{\leq}\; P^{-\min\limits_{\boldsymbol{\mu}\in\mathcal{R}} \mathbf{1}\cdot\boldsymbol{\mu}} \overset{(a)}{\doteq} P^{-\frac{L}{\max_{e\in E}\beta_e}\left(1-\frac{S}{L}r\right)^+}, \tag{37}$$

where (a) follows from the fact that for every $\boldsymbol{\mu} \in \mathcal{R}$, $(L-Sr)^+ \leq \sum_{e\in E}\beta_e\mu_e \leq \max_{e\in E}\beta_e \sum_{e\in E}\mu_e$ which implies that $\sum_{e\in E}\mu_e = \mathbf{1}\cdot\boldsymbol{\mu} \geq \frac{(L-Sr)^+}{\max_{e\in E}\beta_e}$, and on the other hand, defining $\boldsymbol{\mu}^\star$ such that $\boldsymbol{\mu}^\star(\hat{e}) =$

[10] More precisely, with probability greater than $1 - P^{-\delta}$, for any $\delta > 0$.

$\frac{(L-Sr)^{+}}{\beta_{\hat{e}}}$ where $\hat{e} = \underset{e \in E}{\operatorname{argmax}}\ \beta_e$ and otherwise $\boldsymbol{\mu}^{\star}(e) = 0$, we have $\boldsymbol{\mu}^{\star} \in \mathcal{R}$ and $\mathbf{1}\cdot\boldsymbol{\mu}^{\star} = \frac{L}{\max_{e \in E} \beta_e}\left(1 - \frac{S}{L}r\right)^{+}$. (37) completes the proof of Theorem 4. ■

Remark 5- The lower-bound of (6) can also be proved by using the lower-bound of (31) obtained for DMT of the general RS scheme. In order to prove this, one needs to apply the RS scheme with the same path sequence and timing sequence used in the proof of (6) in Theorem 1. Putting $S = L_0 d_G$ and $S \leq l_G L$ in (31) and noting that for all $e \in E$, we have $\beta_e \in \{0, L_0\}$, (6) is easily obtained.

Remark 6- It should be noted that (5) is yet an upper-bound for the DMT of the RS scheme, i.e., even for the case of interfering relays. This is due to the fact that in the proof of (5) the non-interfering relaying assumption is not used. However, by employing the RS scheme with causal-interfering relaying and applying (31), one can find a bigger family of graph topologies that can achieve (5). Such an example is the two-hop relay network studied in Theorem 3. Another example is the case that G is a directed acyclic graph (DAG)[11] and the relays are operating in the full-duplex mode. Here, the argument is similar to that of *Remark 2*. Assume that each $\hat{\mathrm{p}}_{\mathrm{i}}$ is used L_0 times in the path sequence in the form that $\mathrm{p}_{(i-1)L_0+j} \triangleq \hat{\mathrm{p}}_{\mathrm{i}}, 1 \leq j \leq L_0$. Let us modify the timing sequence as $s_{i,j} = i + j - 1 + \sum_{k=1}^{\lceil \frac{i}{L_0} \rceil - 1} \hat{l}_k$ which results in $S = L + \sum_{i=1}^{d_G} l_i$. Here, it is easy to verify that only non-causal interference exists between the signals corresponding to different paths. However, by considering the paths in the reverse order or equivalently reversing the time axis, the paths can be observed with the causal interference. Hence, the result of Theorem 4 is still valid for such paths. Here, knowing that for all $e \in E$, we have $\beta_e \in \{0, L_0\}$ and applying (31), we have $d_{RS}(r) \geq d_G \left(1 - r - \frac{\sum_{i=1}^{d_G} l_i}{L_0 d_G}\right)^{+}$ which achieves (5) for asymptotically large values of L_0. This fact is also observed by [28].

C. Multiple-Access Multiple-Relay Scenario

In this subsection, we generalize the result of Theorem 3 to the multiple-access scenario aided by multiple relay nodes. Here, similar to Theorem 3, we assume that there is no direct link between each transmitter and the receiver. However, no restriction is imposed on the induced subgraph of G on the relay nodes. Assuming having M transmitters, we show that for the rate sequence $r_1 \log(P), r_2 \log(P), \ldots, r_M \log(P)$, in the asymptotic case of $B \to \infty$ (B is the number of sub-blocks), the RS scheme achieves the diversity gain $d_{SM,MAC}(r_1, r_2, \ldots, r_M) = K\left(1 - \sum_{m=1}^{M} r_m\right)^{+}$, which is shown to be optimum due to the cut-set bound on the cutset between the relays and the receiver. Here, the notations are slightly modified compared to the ones used in Theorem 3 to emphasize the fact that multiple signals are transmitted from multiple transmitters. Throughout this subsection and the next one, $\mathbf{x}_m$ and $h_{m,k}$ denote the transmitted vector at

[11] A directed acyclic graph G is a directed graph that has no directed cycles.

the m'th transmitter and the Rayleigh channel coefficient between the m'th transmitter and the k'th relay, respectively. Hence, at the received side of the k'th relay, we have

$$\mathbf{r}_k = \sum_{m=1}^{M} h_{m,k}\mathbf{x}_m + i_{(k)}\mathbf{t}_{(k)} + \mathbf{n}_k, \tag{38}$$

where $\mathbf{x}_m$ is the transmitted vector of the m'th sender. The amplification coefficient at the k'th relay is set to

$$\alpha_k = \min\left\{1, \sqrt{\frac{P}{P\left(\sum_{m=1}^{M}|h_{m,k}|^2 + \left|i_{(k)}\right|^2\right)+1}}\right\}. \tag{39}$$

Here, the RS scheme is applied with the same path sequence and timing sequence as in the case of Theorem 2 and 3. However, it should be mentioned that in the current case, during the slots that the transmitter is supposed to transmit the signal, i.e. in the $s_{i,1}$'th slot, all the transmitters send their signals coherently. Moreover, at the receiver side, after receiving the BK vectors corresponding to the outputs of the BK paths, the destination node decodes the messages $\omega_1, \omega_2, \ldots, \omega_K$ by joint-typical decoding of the received vectors in the corresponding BK slots and the transmitted signal of all the transmitters, i.e., in the same way that joint-typical decoding works in the multiple access setup [10]. Now, we prove the main result of this subsection.

Theorem 5 *Consider a multiple-access channel consisting of M transmitting nodes aided by $K > 1$ half-duplex relays. Assume there is no direct link between the transmitters and the receiver. The RS scheme with the path sequence and timing sequence defined in Theorems 2 and 3 achieves a diversity gain of*

$$d_{RS,MAC}(r_1, r_2, \ldots, r_M) \geq \left[K\left(1 - \sum_{m=1}^{M} r_m\right) - \frac{\sum_{m=1}^{M} r_m}{B}\right]^+, \tag{40}$$

where $r_1, r_2, \ldots, r_M$ are the multiplexing gains corresponding to users $1, 2, \ldots, M$. Moreover, as $B \to \infty$, it achieves the optimum DMT which is $d_{opt,MAC}(r_1, r_2, \ldots, r_M) = K\left(1 - \sum_{m=1}^{M} r_m\right)^+$.

Proof: At the receiver side, we have

$$\begin{aligned}\mathbf{y}_{b,k} &= g_{(k)}\mathbf{t}_{b,k} + \mathbf{z}_{b,k} \\ &= g_{(k)}\alpha_{(k)}\left(\sum_{\substack{1\leq b_1\leq b, 1\leq k_1\leq K \\ b_1K+k_1<bK+k}} p_{b-b_1,k,k_1}\left(\sum_{m=1}^{M} h_{m,k_1}\mathbf{x}_{m,b_1,k_1} + \mathbf{n}_{b_1,k_1}\right)\right) + \mathbf{z}_{b,k},\end{aligned} \tag{41}$$

where p_{b,k,k_1} is defined in the proof of Theorem 3 and $\mathbf{x}_{m,b,k}$ represents the transmitted signal of the m'th sender in the k'th slot of the b'th sub-block. Similar to (27), we have

$$\mathbf{y}(s) = \mathbf{G\Omega F}\left(\sum_{m=1}^{M}\mathbf{H}_m\mathbf{x}_m(s) + \mathbf{n}(s)\right) + \mathbf{z}(s), \tag{42}$$

where $\mathbf{H}_m = \mathbf{I}_B \otimes \textit{diag}\{h_{m,1}, h_{m,2}, \cdots, h_{m,K}\}$, $\mathbf{x}_m(s) = [x_{m,1,1}(s), x_{m,1,2}(s), \cdots, x_{m,B,K}(s)]^T$, and $\mathbf{y}_s, \mathbf{n}_s, \mathbf{z}_s, \mathbf{G}, \mathbf{\Omega}, \mathbf{F}$ are defined in the proof of Theorem 3. Similarly, we observe that the entire channel from each of the transmitters to the receiver acts as a MIMO channel with a lower triangular matrix of size $BK \times BK$.

Here, the outage event occurs whenever there exists a subset $\mathcal{S} \subseteq \{1, 2, \ldots, M\}$ of the transmitters such that

$$I\left(\mathbf{x}_{\mathcal{S}}(s); \mathbf{y}(s) | \mathbf{x}_{\mathcal{S}^c}(s)\right) \leq (BK+1)\left(\sum_{m\in\mathcal{S}} r_m\right)\log(P). \tag{43}$$

This event is equivalent to

$$\log\left|\mathbf{I}_{BK} + P\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right| \leq (BK+1)\left(\sum_{m\in\mathcal{S}} r_m\right)\log(P). \tag{44}$$

where $\mathbf{P}_n$ is defined in the proof of Theorem 3, $\mathbf{H}_T = \mathbf{G\Omega F H}_{\mathcal{S}}$, and

$$\mathbf{H}_{\mathcal{S}} = \mathbf{I}_B \otimes \textit{diag}\left\{\sqrt{\sum_{m\in\mathcal{S}} |h_{m,1}|^2}, \sqrt{\sum_{m\in\mathcal{S}} |h_{m,2}|^2}, \cdots, \sqrt{\sum_{m\in\mathcal{S}} |h_{m,K}|^2}\right\}. \tag{45}$$

Defining such an event as $\mathcal{E}_{\mathcal{S}}$ and the outage event as $\mathcal{E}$, we have

$$\begin{aligned}\mathbb{P}\{\mathcal{E}\} &= \mathbb{P}\left\{\bigcup_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathcal{E}_S\right\} \\ &\leq \sum_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathbb{P}\{\mathcal{E}_S\} \\ &\leq (2^M-1)\max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathbb{P}\{\mathcal{E}_S\} \\ &\doteq \max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathbb{P}\{\mathcal{E}_S\}.\end{aligned} \tag{46}$$

Hence, it is sufficient to upper-bound $\mathbb{P}\{\mathcal{E}_S\}$ for all $\mathcal{S}$.

Defining $\hat{\mathbf{H}}_{\mathcal{S}} = \mathbf{I}_B \otimes \textit{diag}\{\max_{m\in\mathcal{S}}|h_{m,1}|, \max_{m\in\mathcal{S}}|h_{m,2}|, \cdots, \max_{m\in\mathcal{S}}|h_{m,K}|\}$, we have $\hat{\mathbf{H}}_{\mathcal{S}}\hat{\mathbf{H}}_{\mathcal{S}}^H \preccurlyeq \mathbf{H}_{\mathcal{S}}\mathbf{H}_{\mathcal{S}}^H$. Therefore,

$$\begin{aligned}\mathbb{P}\{\mathcal{E}_S\} &\leq \mathbb{P}\left\{\log\left|\mathbf{I}_{BK} + P\mathbf{G\Omega F}\hat{\mathbf{H}}_{\mathcal{S}}\hat{\mathbf{H}}_{\mathcal{S}}^H\mathbf{F}^H\mathbf{\Omega}^H\mathbf{G}^H\mathbf{P}_n^{-1}\right| \leq (BK+1)\left(\sum_{m\in\mathcal{S}} r_m\right)\log(P)\right\} \\ &\triangleq \mathbb{P}\left\{\hat{\mathcal{E}}_{\mathcal{S}}\right\}.\end{aligned} \tag{47}$$

Assume $\max_{m\in\mathcal{S}}|h_{m,k}|^2 = P^{-\mu_k}$, and $|g_k|^2 = P^{-\nu_k}$, $|i_k|^2 = P^{-\omega_k}$, and $\mathcal{R}$ as the region in $\mathbb{R}^{3K}$ that defines the outage event $\hat{\mathcal{E}}_{\mathcal{S}}$ in terms of the vector $[\boldsymbol{\mu}^T, \boldsymbol{\nu}^T, \boldsymbol{\omega}^T]^T$. Similar to the proof of Theorem 3, we have $\mathbb{P}\{\mathcal{R}\} \doteq \mathbb{P}\{\mathcal{R}_+\}$ where $\mathcal{R}_+ = \mathcal{R}\bigcap\mathbb{R}_+^{3K}$. Rewriting the equation series of (29), we have

$$\begin{aligned}\mathbb{P}\left\{\hat{\mathcal{E}}_{\mathcal{S}}\right\} &\dot{\leq} \mathbb{P}\left\{-BK\frac{\log\left[3\left(B^2K^2+1\right)\right]}{\log(P)} + BK\left(1 - \sum_{m\in\mathcal{S}} r_m\right) - \sum_{m\in\mathcal{S}} r_m \leq B\sum_{k=1}^{K}(\mu_k+\nu_k),\right. \\ &\qquad\left. \mu_k, \nu_k, \omega_k \geq 0\right\}.\end{aligned} \tag{48}$$

On the other hand, as $\{h_{m,k}\}$'s are independent random variables, we conclude that for $\boldsymbol{\mu}^0, \boldsymbol{\nu}^0 \geq \mathbf{0}$, we have $\mathbb{P}\left\{\boldsymbol{\mu} \geq \boldsymbol{\mu}^0, \boldsymbol{\nu} \geq \boldsymbol{\nu}^0\right\} \doteq P^{-\mathbf{1}\cdot\left(|\mathcal{S}|\boldsymbol{\mu}^0+\boldsymbol{\nu}^0\right)}$. Similar to the proof of Theorem 3, by computing the derivative with respect to $\boldsymbol{\mu}, \boldsymbol{\nu}$, we have $f_{\boldsymbol{\mu},\boldsymbol{\nu}}(\boldsymbol{\mu}, \boldsymbol{\nu}) \doteq P^{-\mathbf{1}\cdot(|\mathcal{S}|\boldsymbol{\mu}+\boldsymbol{\nu})}$. Defining $l_0 \triangleq -\frac{\log\left[3\left(B^2K^2+1\right)\right]}{\log(P)} + \left(1 - \sum_{m\in\mathcal{S}} r_m\right) - \frac{\sum_{m\in\mathcal{S}} r_m}{BK}$, the region $\hat{\mathcal{R}}$ as $\hat{\mathcal{R}} \triangleq \left\{\boldsymbol{\mu}, \boldsymbol{\nu} \geq \mathbf{0}, \frac{1}{K}\mathbf{1}\cdot(\boldsymbol{\mu}+\boldsymbol{\nu}) \geq l_0\right\}$, the cube $\mathcal{I}$ as $\mathcal{I} \triangleq [0, Kl_0]^{2K}$, and for $1 \leq i \leq 2K$, $\mathcal{I}_i^c = [0,\infty)^{i-1} \times [Kl_0, \infty) \times [0,\infty)^{2K-i}$, we have

$$\begin{aligned}
\mathbb{P}\left\{\hat{\mathcal{E}}_{\mathcal{S}}\right\} &\overset{(a)}{\dot{\leq}} \mathbb{P}\{\hat{\mathcal{R}}\} \\
&\leq \int_{\hat{\mathcal{R}}\bigcap\mathcal{I}} f_{\boldsymbol{\mu},\boldsymbol{\nu}}(\boldsymbol{\mu},\boldsymbol{\nu})\, d\boldsymbol{\mu} d\boldsymbol{\nu} + \sum_{i=1}^{2K} \mathbb{P}\left\{[\boldsymbol{\mu}^T, \boldsymbol{\nu}^T]^T \in \hat{\mathcal{R}} \cap \mathcal{I}_i^c\right\} \\
&\dot{\leq} \operatorname{vol}(\hat{\mathcal{R}} \cap \mathcal{I}) P^{-\min\limits_{[\boldsymbol{\mu},\boldsymbol{\nu}]\in\hat{\mathcal{R}}\bigcap\mathcal{I}} \mathbf{1}\cdot(|\mathcal{S}|\,\boldsymbol{\mu}+\boldsymbol{\nu})} + 2KP^{-Kl_0} \\
&\overset{(b)}{\doteq} P^{-Kl_0} \\
&\doteq P^{-\left[K\left(1-\sum_{m\in\mathcal{S}} r_m\right) - \frac{\sum_{m\in\mathcal{S}} r_m}{B}\right]}.
\end{aligned} \tag{49}$$

Here, (a) follows from (48) and (b) follows from the fact that $\hat{\mathcal{R}}\bigcap\mathcal{I}$ is a bounded region whose volume is independent of P and the fact that $\min_{[\boldsymbol{\mu},\boldsymbol{\nu}]\in\hat{\mathcal{R}}\bigcap\mathcal{I}} \mathbf{1}\cdot(|\mathcal{S}|\,\boldsymbol{\mu}+\boldsymbol{\nu}) = Kl_0$, which is achieved by having $\boldsymbol{\mu} = \mathbf{0}$. Comparing (46), (47) and (49), we observe

$$\mathbb{P}\left\{\mathcal{E}\right\} \dot{\leq} \max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathbb{P}\left\{\mathcal{E}_{\mathcal{S}}\right\} \dot{\leq} \max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathbb{P}\left\{\hat{\mathcal{E}}_{\mathcal{S}}\right\} \dot{\leq} P^{-\left[K\left(1-\sum_{m=1}^{M} r_m\right) - \frac{\sum_{m=1}^{M} r_m}{B}\right]}. \tag{50}$$

Next, we prove that $K\left(1 - \sum_{m=1}^{M} r_m\right)^+$ is an upper-bound on the diversity gain of the system corresponding to the sequence of rates $r_1, r_2, \ldots, r_M$. We have

$$\mathbb{P}\left\{\mathcal{E}\right\} \geq \mathbb{P}\left\{\max_{p(\mathbf{t}_1,\mathbf{t}_2,\ldots,\mathbf{t}_K)} I\left(\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_K; \mathbf{y}\right) \leq \left(\sum_{m=1}^{M} r_m\right)\log(P)\right\} \overset{(a)}{\doteq} P^{-K\left(1-\sum_{m=1}^{M} r_m\right)^+}. \tag{51}$$

Here, (a) follows from the DMT of the point-to-point MISO channel proved in [9]. This completes the proof. ■

Remark 7- The argument of Theorem 5 is valid for the general case in which any arbitrary set of relay pairs are non-interfering.

Remark 8- In the *symmetric* situation for which the multiplexing gains of all the users are equal to say r, the lower-bound in (40) takes a simple form. First, we observe that the maximum multiplexing gain which is simultaneously achievable by all the users is $\frac{1}{M} \cdot \frac{BK}{BK+1}$. Noting that no signal is sent to the receiver in $\frac{1}{BK+1}$ portion of the time, we observe that the RS scheme achieves the maximum possible symmetric multiplexing gain for all the users. Moreover, from (40), we observe that the RS scheme achieves the maximum diversity gain of K for any finite value of B, which turns out to be tight as well. Finally, the lower-bound on the DMT of the RS scheme is simplified to $\left[K\left(1 - Mr\right) - \frac{Mr}{B}\right]^+$ for the *symmetric* situation.

D. Multiple-Access Single Relay Scenario

As we observe, the arguments of Theorems 2, 3 and 5 concerning DMT optimality of the RS scheme are valid for the scenario of having multiple relays ($K > 1$). Indeed, for the single relay scenario, the RS scheme is reduced to the simple amplify-and-forward relaying in which the relay listens to the transmitter in the first half of the frame and transmits the amplified version of the received signal in the second half. However, like the case of non-interfering relays studied in [14], the DMT optimality arguments are no longer valid. On the other hand, we show that the DDF scheme achieves the optimum DMT for this scenario.

Theorem 6 *Consider a multiple-access channel consisting of M transmitting nodes aided by a single half-duplex relay. Assume that all the network nodes are equipped with a single antenna and there is no direct link between the transmitters and the receiver. The amplify-and-forward scheme achieves the following DMT*

$$d_{AF,MAC}(r_1, r_2, \ldots, r_M) = \left(1 - 2\sum_{m=1}^{M} r_m\right)^+. \tag{52}$$

However, the optimum DMT of the network is

$$d_{MAC}(r_1, r_2, \ldots, r_M) = \left(1 - \frac{\sum_{m=1}^{M} r_m}{1 - \sum_{m=1}^{M} r_m}\right)^+, \tag{53}$$

which is achievable by the DDF scheme of [6].

Proof: First, we show that the DMT of the AF scheme follows (52). At the receiver side, we have

$$\mathbf{y} = g\alpha\left(\sum_{m=1}^{M} h_m \mathbf{x}_m + \mathbf{n}\right) + \mathbf{z}, \tag{54}$$

where h_m is the channel gain between the m'th transmitter and the relay, g is the down-link channel gain, and $\alpha = \sqrt{\frac{P}{P\sum_{m=1}^{M}|h_m|^2+1}}$ is the amplification coefficient. Defining the outage event $\mathcal{E}_{\mathcal{S}}$ for a set

$\mathcal{S} \subseteq \{1, 2, \ldots, M\}$, similar to the case of Theorem 5, we have

$$\begin{aligned}
\mathbb{P}\left\{\mathcal{E}_{\mathcal{S}}\right\} &= \mathbb{P}\left\{I\left(\mathbf{x}_{\mathcal{S}};\mathbf{y}|\mathbf{x}_{\mathcal{S}^c}\right) < 2\left(\sum_{m\in\mathcal{S}} r_m\right)\log(P)\right\} \\
&= \mathbb{P}\left\{\log\left(1 + P\left(\sum_{m\in\mathcal{S}} |h_m|^2\right)|g|^2|\alpha|^2\left(1+|g|^2|\alpha|^2\right)^{-1}\right) < 2\left(\sum_{m\in\mathcal{S}} r_m\right)\log(P)\right\} \\
&\doteq \mathbb{P}\left\{\left(\sum_{m\in\mathcal{S}} |h_m|^2\right)|g|^2|\alpha|^2 \min\left\{1, \frac{1}{|g|^2|\alpha|^2}\right\} \leq P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} \\
&\overset{(a)}{\doteq} \mathbb{P}\left\{\sum_{m\in\mathcal{S}} |h_m|^2 \leq P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} + \\
&\quad\ \mathbb{P}\left\{\left(\sum_{m\in\mathcal{S}} |h_m|^2\right)|g|^2|\alpha|^2 \leq P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} \\
&\overset{(b)}{\doteq} \mathbb{P}\left\{\sum_{m\in\mathcal{S}} |h_m|^2 \leq P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} + \\
&\quad\ \mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}} |h_m|^2\right)\min\left\{P, \frac{1}{\sum_{m=1}^{M}|h_m|^2}\right\} \leq P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} \\
&\overset{(a)}{\doteq} \mathbb{P}\left\{\sum_{m\in\mathcal{S}} |h_m|^2 \leq P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} + \\
&\quad\ \mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}} |h_m|^2\right) \leq P^{-2\left(1-\sum_{m\in\mathcal{S}} r_m\right)}\right\} + \\
&\quad\ \mathbb{P}\left\{\frac{|g|^2\sum_{m\in\mathcal{S}}|h_m|^2}{\sum_{m=1}^{M}|h_m|^2} \leq P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\}. \qquad (55)
\end{aligned}$$

In the above equation, (a) comes from the fact that $\mathbb{P}\{\min(X,Y) \leq z\} = \mathbb{P}\left\{(X \leq z)\bigcup(Y \leq z)\right\} \doteq \mathbb{P}\{X \leq z\} + \mathbb{P}\{Y \leq z\}$. (b) follows from the fact that $|\alpha|^2$ can be asymptotically written as $\min\left\{P, \frac{1}{\sum_{m=1}^{M}|h_m|^2}\right\}$.

Since $\{|h_m|^2\}_{m=1}^{M}$ are i.i.d. random variables with exponential distribution, it follows that $\sum_{m\in\mathcal{S}}|h_m|^2$ has Chi-square distribution with $2|\mathcal{S}|$ degrees of freedom, which implies that

$$\mathbb{P}\left\{\sum_{m\in\mathcal{S}} |h_m|^2 \leq P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} \doteq P^{-|\mathcal{S}|\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}. \qquad (56)$$

To compute the second term in (55), defining $\epsilon_1 \triangleq P^{-2\left(1-\sum_{m\in\mathcal{S}} r_m\right)}$, we have

$$\begin{aligned}
\mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}} |h_m|^2\right) \leq \epsilon_1\right\} &\overset{(a)}{\dot{\geq}} \mathbb{P}\left\{|g|^2 \leq \epsilon_1\right\} \\
&\doteq \epsilon_1, \qquad (57)
\end{aligned}$$

where (a) follows from the fact that as $\sum_{m\in\mathcal{S}}|h_m|^2$ has Chi-square distribution, we have $\sum_{m\in\mathcal{S}}|h_m|^2 \dot{\leq} 1$ with probability one (more precisely, with a probability greater than $1 - P^{-\delta}$ for every $\delta > 0$). On the

other hand, we have

$$\mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}}|h_m|^2\right)\leq\epsilon_1\right\} \leq \mathbb{P}\left\{|g|^2|h_m|^2\leq\epsilon_1\right\} \doteq \epsilon_1. \tag{58}$$

Putting (57) and (58) together, we have

$$\mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}}|h_m|^2\right)\leq\epsilon_1\right\}\doteq\epsilon_1. \tag{59}$$

Now, to compute the third term in (55), defining $\epsilon_2\triangleq P^{-\left(1-2\sum_{m\in\mathcal{S}}r_m\right)}$, we observe

$$\epsilon_2\doteq\mathbb{P}\left\{|g|^2\leq\epsilon_2\right\}\leq\mathbb{P}\left\{|g|^2\frac{\sum_{m\in\mathcal{S}}|h_m|^2}{\sum_{m=1}^{M}|h_m|^2}\leq\epsilon_2\right\}\overset{(a)}{\dot{\leq}}\mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}}|h_m|^2\right)\leq\epsilon_2\right\}\overset{(b)}{\doteq}\epsilon_2.$$

Here, (a) follows from the fact that with probability one, we have $\sum_{m=1}^{M}|h_m|^2\dot{\leq}1$ and (b) follows from (59). As a result

$$\mathbb{P}\left\{|g|^2\frac{\sum_{m\in\mathcal{S}}|h_m|^2}{\sum_{m=1}^{M}|h_m|^2}\leq\epsilon_2\right\}\doteq\epsilon_2 \tag{60}$$

From (56), (59), and (60), we have

$$\mathbb{P}\left\{\mathcal{E}_{\mathcal{S}}\right\}\doteq P^{-|\mathcal{S}|\left(1-2\sum_{m\in\mathcal{S}}r_m\right)^+}+P^{-2\left(1-\sum_{m\in\mathcal{S}}r_m\right)^+}+P^{-\left(1-2\sum_{m\in\mathcal{S}}r_m\right)^+}\doteq P^{-\left(1-2\sum_{m\in\mathcal{S}}r_m\right)^+}. \tag{61}$$

Observing (61) and applying the argument of (46), we have

$$\mathbb{P}\left\{\mathcal{E}\right\}\doteq\max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}}\mathbb{P}\left\{\mathcal{E}_{\mathcal{S}}\right\}\doteq P^{-\left(1-2\sum_{m=1}^{M}r_m\right)^+}. \tag{62}$$

This completes the proof for the AF scheme. Now, to compute the DMT of the DDF scheme, let us assume that the relay listens to the transmitted signal for the l portion of the time until it can decode it perfectly. Hence, we have

$$l=\min\left\{1,\max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}}\frac{\left(\sum_{m\in\mathcal{S}}r_m\right)\log(P)}{\log\left(1+\left(\sum_{m\in\mathcal{S}}|h_m|^2\right)P\right)}\right\}. \tag{63}$$

The outage event occurs whenever the relay can not transmit the re-encoded information in the remaining portion of the time. Hence, we have

$$\mathbb{P}\left\{\mathcal{E}\right\}\doteq\mathbb{P}\left\{(1-l)\log\left(1+|g|^2P\right)<\left(\sum_{m=1}^{M}r_m\right)\log(P)\right\}. \tag{64}$$

Assuming $|h_m|^2=P^{-\mu_m}$ and $|g|^2=P^{-\nu}$, at high SNR, we have

$$l\approx\min\left\{1,\max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}}\frac{\sum_{m\in\mathcal{S}}r_m}{1-\min_{m\in\mathcal{S}}\mu_m}\right\}. \tag{65}$$

Equivalently, an outage event occurs whenever

$$\left(1-\max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}}\frac{\sum_{m\in\mathcal{S}}r_m}{1-\min_{m\in\mathcal{S}}\mu_m}\right)(1-\nu)<\sum_{m=1}^{M}r_m. \tag{66}$$

In order to find the probability of the outage event, we first find an upper-bound on the outage probability and then, we show that this upper-bound is indeed tight. Defining $R = \sum_{m=1}^{M} r_M$ and $\mu = \sum_{m=1}^{M} \mu_m$, we have

$$R \overset{(a)}{>} \left(1 - \frac{\sum_{m \in \mathcal{S}_0} r_m}{1 - \min_{m \in \mathcal{S}_0} \mu_m}\right)(1-\nu) > \left(1 - \frac{R}{1-\mu}\right)(1-\nu)\,. \tag{67}$$

Here, (a) follows from (66). Equivalently,

$$R \overset{(a)}{>} \frac{(1-\mu)(1-\nu)}{(1-\mu)+(1-\nu)} > \frac{1-\mu-\nu}{(1-\mu)+(1-\nu)}, \tag{68}$$

where (a) follows from (67). It can be easily checked that (68) is equivalent to

$$R > (1-R)(1-\mu-\nu). \tag{69}$$

In other words, any vector point $[\mu_1, \mu_2, \ldots, \mu_M, \nu]$ in the outage region $\mathcal{R}$, i.e., the region that satisfies (66), also satisfies (69). As a result, defining $\mathcal{R}'$ as the region defined by (69), we have

$$\mathbb{P}\{\mathcal{E}\} \;\; \leq \;\; \mathbb{P}\{\boldsymbol{\pi} \in \mathcal{R}'\}, \tag{70}$$

where $\boldsymbol{\pi} \triangleq [\mu_1, \mu_2, \ldots, \mu_M, \nu]$. Similar to the approach used in the proofs of Theorems 3 and 5, $\mathbb{P}\{\boldsymbol{\pi} \in \mathcal{R}'\}$ can be computed as

$$\mathbb{P}\{\boldsymbol{\pi} \in \mathcal{R}'\} \doteq P^{-\frac{R}{1-R}}. \tag{71}$$

Hence,

$$\mathbb{P}\{\mathcal{E}\} \dot{\leq} P^{-\frac{R}{1-R}}. \tag{72}$$

For lower-bounding the outage probability, we note that all the vectors $[\mu_1, \cdots, \mu_M, \nu]$ for which $\mu_m > 0, m = 1, \cdots, M$ and $\nu > \frac{R}{1-R}$, lie in the outage region defined in (66). In other words,

$$\begin{aligned} \mathbb{P}\{\mathcal{E}\} \;\; &\geq \;\; \mathbb{P}\left\{\boldsymbol{\pi} > \left[0, \cdots, 0, \frac{R}{1-R}\right]\right\} \\ &\doteq \;\; P^{-\frac{R}{1-R}}. \end{aligned} \tag{73}$$

Combining (72) and (73) yields

$$\begin{aligned} \mathbb{P}\{\mathcal{E}\} \;\; &\doteq \;\; P^{-\frac{R}{1-R}} \\ &= \;\; P^{-\frac{\sum_{m=1}^{M} r_m}{1-\sum_{m=1}^{M} r_m}}, \end{aligned} \tag{74}$$

which completes the proof for the DMT analysis of the DDF scheme.

Next, we prove that the DDF scheme achieves the optimum DMT. As the channel from the transmitters to the receiver is a degraded version of the channel between the transmitters and the relay, similar to the argument of [31] for the case of single-source single-relay, we can easily show that the decode-forward

strategy achieves the capacity of the network for each realization of the channels. Now, consider the realization in which for all m we have, $|h_m|^2 \leq \frac{1}{M}$. As we know, $\mathbb{P}\left\{\forall m : |h_m|^2 \leq \frac{1}{M}\right\} \doteq 1$. Let us assume in the optimum decode-and-forward strategy, the relay spends l portion of the time for listening to the transmitter. According to the Fano's inequality [10], to make the probability of error in decoding the transmitters' message at the relay side approach zero, we should have $l \log\left(1 + \frac{P}{l}\sum_{m=1}^{M}|h_m|^2\right) \geq \left(\sum_{m=1}^{M} r_m\right)\log(P)$. Accordingly, we should have $l \geq \sum_{m=1}^{M} r_m$. On the other hand, in order that the receiver can decode the relay's message with a vanishing probability of error in the remaining portion of the time, we should have $(1-l)\log\left(1 + \frac{P}{1-l}|g|^2\right) \geq \sum_{m=1}^{M} r_m \log(P)$. Hence, we have $\mathbb{P}\left\{\mathcal{E}\right\} \geq \mathbb{P}\left\{|g|^2 \leq cP^{-\left(1-\frac{\sum_{m=1}^{M} r_m}{1-\sum_{m=1}^{M} r_m}\right)}, \forall m : |h_m|^2 \leq \frac{1}{M}\right\} \doteq P^{-\left(1-\frac{\sum_{m=1}^{M} r_m}{1-\sum_{m=1}^{M} r_m}\right)^+}$, for a constant c. This completes the proof. ∎

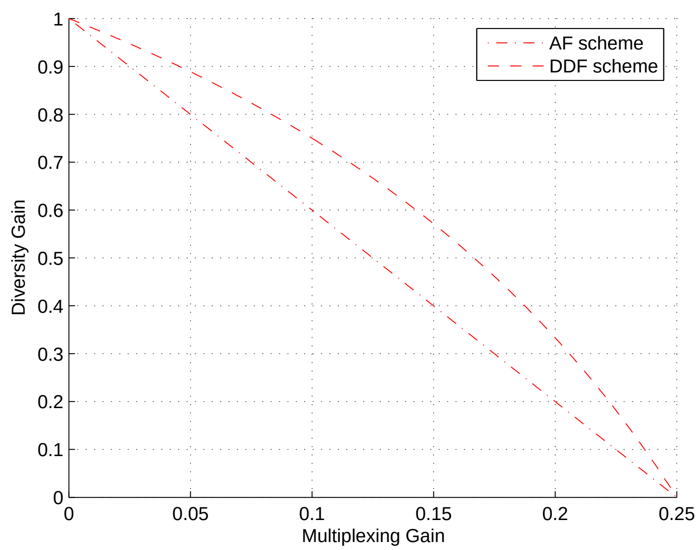

Fig. 4. Diversity-Multiplexing Tradeoff of AF scheme versus the optimum and DDF scheme for multiple access single relay channel consisting of $M = 2$ transmitters assuming *symmetric* transmission, i.e. $r_1 = r_2 = r$.

Figure 4 shows DMT of the AF scheme and the DDF scheme for multiple access single relay setup consisting of $M = 2$ transmitters assuming *symmetric* situation, i.e. $r_1 = r_2 = r$. As can be observed in this figure, although the AF scheme achieves the maximum multiplexing gain and maximum diversity gain, it does not achieve the optimum DMT in any other points of the tradeoff region.

V. Maximum Diversity Achievability Proof in General Multi-Hop Multiple-Antenna Scenario

In this section, we consider our proposed RS scheme and prove that it achieves the maximum diversity gain between two end-points in a general multiple-antenna multi-hop network (no additional constraints

imposed). However, in this general scenario, it can not achieve the optimum DMT. Indeed, we show that in order to achieve the optimum DMT, in some scenarios, multiple interfering nodes have to transmit together during the same slot.

Theorem 7 *Consider a relay network with the connectivity graph $G = (V, E)$ and K relays, in which each two adjacent nodes are connected through a Rayleigh-fading channel. Assume that all the network nodes are equipped with multiple antennas. Then, by properly choosing the path sequence, the proposed RS scheme achieves the maximum diversity gain of the network which is equal to*

$$d_G = \min_{\mathcal{S}} w_G(\mathcal{S}), \tag{75}$$

where $\mathcal{S}$ is a cut-set on G.

Proof: First, we show that d_G is indeed an upper-bound on the diversity-gain of the network. To show this, we do not consider the half-duplex nature of the relay nodes and assume that they operate in full-duplex mode. Consider a cut-set $\mathcal{S}$ on G. We have

$$\begin{aligned} \mathbb{P}\left\{\mathcal{E}\right\} &\overset{(a)}{\dot{\geq}} \mathbb{P}\left\{I\left(X\left(\mathcal{S}\right); Y\left(\mathcal{S}^c\right) | X\left(\mathcal{S}^c\right)\right) < R\right\} \\ &\overset{(b)}{=} \mathbb{P}\left\{\sum_{k \in \mathcal{S}^c} I\left(X\left(\mathcal{S}\right); Y_k | Y\left(\mathcal{S}^c / \left\{1, 2, \ldots, k\right\}\right), X\left(\mathcal{S}^c\right)\right) < R\right\} \\ &\overset{(c)}{\geq} \prod_{k \in \mathcal{S}^c} \mathbb{P}\left\{I\left(X\left(\mathcal{S}\right); Y_k | X\left(\mathcal{S}^c\right)\right) < \frac{R}{|\mathcal{S}^c|}\right\} \\ &\overset{(d)}{\doteq} \prod_{k \in \mathcal{S}^c} P^{-|\{e \in E | k \in e, e \cap \mathcal{S} \neq \emptyset\}|} \\ &\doteq P^{-w_G(\mathcal{S})}, \end{aligned} \tag{76}$$

where R is the target rate which does not scale with P (i.e., $r = 0$). Here, (a) follows from the cut-set bound theorem [10] and the fact that for the rates above the capacity, the error probability approaches one (according to Fano's inequality [10]), (b) follows from the chain rule on the mutual information [10], (c) follows from the facts that i) $\left(Y_k, X\left(\left\{0, 1, \ldots, K+1\right\}\right), Y\left(\mathcal{S}^c / \left\{1, 2, \ldots, k\right\}\right)\right)$ form a Markov chain [10] and as a result, $I\left(X\left(\mathcal{S}\right); Y_k | Y\left(\mathcal{S}^c / \left\{1, 2, \ldots, k\right\}\right), X\left(\mathcal{S}^c\right)\right) \leq I\left(X\left(\mathcal{S}\right); Y_k | X\left(\mathcal{S}^c\right)\right)$, and ii) $I\left(X\left(\mathcal{S}\right); Y_k | X\left(\mathcal{S}^c\right)\right)$ depends only on the channel matrices between $X\left(\mathcal{S}\right)$ and Y_k and as all the channels in the network are independent of each other, it follows that the events

$$\left\{I\left(X\left(\mathcal{S}\right); Y_k | X\left(\mathcal{S}^c\right)\right) < \frac{R}{|\mathcal{S}^c|}\right\}_{k \in \mathcal{S}^c}$$

are mutually independent, and finally (d) follows from the diversity gain of the MISO channel. Considering all possible cut-sets on G and using (76), we have

$$\mathbb{P}\left\{\mathcal{E}\right\} \dot{\geq} P^{-\min_{\mathcal{S}} w_G(\mathcal{S})}. \tag{77}$$

Now, we prove that this bound is indeed achievable by the RS scheme. First, we provide the path sequence needed to achieve the maximum diversity gain. Consider the graph $\hat{G} = (V, E, w)$ with the same set of vertices and edges as the graph G and the weight function w on the edges as $w_{\{a,b\}} = N_a N_b$. Consider the maximum-flow algorithm [29] on $\hat{G}$ from the source node 0 to the sink node $K+1$. Since the weight function is integer over the edges, according to the Ford-Fulkerson Theorem [29], one can achieve the maximum flow which is equal to the minimum cut of $\hat{G}$ or d_G by the union of elements of a sequence $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_{d_G})$ of paths ($L = d_G$). We show that this family of paths are sufficient to achieve the optimum diversity. Here, we do not consider the problem of selecting the path timing sequence $\{s_{i,j}\}$. We just assume that a timing sequence $\{s_{i,j}\}$ with the 4 requirements defined in the third section exists. However, it should be noted that as we consider the maximum diversity throughout the theorem, we are not concerned with $\frac{S}{L}$. Hence, we can select the path timing sequence such that no two paths cause interference on each other.

Noting that the received signal at each node is multiplied by a random isotropically distributed unitary matrix, at the receiver side we have

$$\begin{aligned} \mathbf{y}_{K+1,i} &= \mathbf{H}_{K+1,\mathrm{p}_i(l_i-1)}\alpha_{i,l_i-1}\mathbf{U}_{i,l_i-1}\mathbf{H}_{\mathrm{p}_i(l_i-1),\mathrm{p}_i(l_i-2)}\alpha_{i,l_i-2}\mathbf{U}_{i,l_i-2}\cdots\alpha_{i,1}\mathbf{U}_{i,1}\mathbf{H}_{\mathrm{p}_i(1),0}\mathbf{x}_{0,i} + \\ &\quad \sum_{j<i}\mathbf{X}_{i,j}\mathbf{x}_{0,j} + \sum_{j\le i, m\le l_j}\mathbf{Q}_{i,j,m}\mathbf{n}_{j,m}. \end{aligned} \tag{78}$$

Here, $\mathbf{x}_{0,i}$ is the vector transmitted at the transmitter side during the $s_{i,1}$'th slot as the input for the i'th path, $\mathbf{y}_{K+1,i}$ is the vector received at the receiver side during the s_{i,l_i}'th slot as the output for i'th path, $\mathbf{U}_{i,j}$ denotes the multiplied unitary matrix at the $\mathrm{p}_i(j)$'th node of the ith path, $\mathbf{X}_{i,j}$ is the interference matrix which relates the input of the j'th path ($j < i$) to the output of the i'th path, $\mathbf{n}_{j,m}$ is the noise vector during the $s_{j,m}$'th slot at the $\mathrm{p}_j(m)$'th node of the network, and finally, $\mathbf{Q}_{i,k,m}$ is the matrix which relates $\mathbf{n}_{k,m}$ to $\mathbf{y}_{K+1,i}$. Notice that as the timing sequence satisfies the noncausal interference assumption, the summation terms in (78) do not exceed i. Defining $\mathbf{x}(s) = \left[\mathbf{x}_{0,1}^T(s)\,\mathbf{x}_{0,2}^T(s)\cdots\mathbf{x}_{0,L}^T(s)\right]^T$, $\mathbf{y}(s) = \left[\mathbf{y}_{K+1,1}^T(s)\,\mathbf{y}_{K+1,2}^T(s)\cdots\mathbf{y}_{K+1,L}^T(s)\right]^T$, and $\mathbf{n}(s) = \left[\mathbf{n}_{1,1}^T(s)\,\mathbf{n}_{1,2}^T(s)\cdots\mathbf{n}_{L,l_L}^T(s)\right]^T$, we have the following equivalent block lower-triangular matrix between the end nodes

$$\mathbf{y}(s) = \mathbf{H}_T\mathbf{x}(s) + \mathbf{Q}\mathbf{n}(s). \tag{79}$$

Here,

$$\mathbf{H}_T = \begin{pmatrix} \mathbf{X}_{1,1} & \mathbf{0} & \mathbf{0} & \ldots \\ \mathbf{X}_{2,1} & \mathbf{X}_{2,2} & \mathbf{0} & \ldots \\ \vdots & \vdots & \vdots & \ddots \\ \mathbf{X}_{L,1} & \mathbf{X}_{L,2} & \ldots & \mathbf{X}_{L,L} \end{pmatrix}, \tag{80}$$

where $\mathbf{X}_{i,i} = \mathbf{H}_{K+1,\mathrm{p}_i(l_i-1)}\alpha_{i,l_i-1}\mathbf{U}_{i,l_i-1}\mathbf{H}_{\mathrm{p}_i(l_i-1),\mathrm{p}_i(l_i-2)}\alpha_{i,l_i-2}\mathbf{U}_{i,l_i-2}\cdots\alpha_{i,1}\mathbf{U}_{i,1}\mathbf{H}_{\mathrm{p}_i(1),0}$, and

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{1,1,1} & \dots & \mathbf{Q}_{1,1,l_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{Q}_{2,1,1} & \dots & \mathbf{Q}_{2,1,l_1} & \dots & \mathbf{Q}_{2,2,l_2} & \mathbf{0} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \mathbf{Q}_{L,1,1} & \mathbf{Q}_{L,1,2} & \dots & \dots & \dots & \mathbf{Q}_{L,L,l_L-1} & \mathbf{Q}_{L,L,l_L} \end{pmatrix}. \tag{81}$$

Having (79), the outage probability can be written as

$$\mathbb{P}\{\mathcal{E}\} \;=\; \mathbb{P}\left\{\left|\mathbf{I}_L + P\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right| < 2^{SR}\right\}, \tag{82}$$

where $\mathbf{P}_n = \mathbf{Q}\mathbf{Q}^H$. First, similar to the proof of theorem 3, we can show that $\alpha_{i,j} \doteq 1$ with probability 1[12], and also show that there exists a constant c which depends just on the topology of graph G and the path sequence such that $\mathbf{P}_n \preccurlyeq c\mathbf{I}_L$. Assume that for each $\{a,b\} \in E$, $\lambda_{\max}\left(\mathbf{H}_{a,b}\right) = P^{-\mu_{\{a,b\}}}$, where $\lambda_{\max}\left(\mathbf{A}\right)$ denotes the greatest eigenvalue of $\mathbf{A}\mathbf{A}^H$. Also, assume that

$$\gamma_{i,j} \triangleq \left|\mathbf{v}_{r,\max}^H\left(\mathbf{H}_{\{\mathrm{p}_i(j+1),\mathrm{p}_i(j)\}}\right)\mathbf{U}_{i,j}\mathbf{v}_{l,\max}\left(\mathbf{H}_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\mathbf{U}_{i,j-1}\mathbf{H}_{\{\mathrm{p}_i(j-1),\mathrm{p}_i(j-2)\}}\dots\mathbf{H}_{\{\mathrm{p}_i(1),0\}}\right)\right|^2 = P^{-\nu_{i,j}}, \tag{83}$$

where $\mathbf{v}_{l,\max}\left(\mathbf{A}\right)$ and $\mathbf{v}_{r,\max}\left(\mathbf{A}\right)$ denote the left and the right eigenvectors of $\mathbf{A}$ corresponding to $\lambda_{\max}\left(\mathbf{A}\right)$, respectively. The outage probability can be upper-bounded as

$$\begin{aligned} \mathbb{P}\{\mathcal{E}\} &\overset{(a)}{\leq} \mathbb{P}\left\{\lambda_{\max}\left(\left(\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right)^{\frac{1}{2}}\right) \leq \left(2^{SR}-1\right)P^{-1}\right\} \\ &\overset{(b)}{\leq} \mathbb{P}\left\{\lambda_{\max}\left(\mathbf{H}_T\right) \leq c\left(2^{SR}-1\right)P^{-1}\right\} \\ &\overset{(c)}{\leq} \mathbb{P}\left\{\bigcap_{i=1}^{L}\left(\lambda_{\max}\left(\mathbf{X}_{i,i}\right) \leq c\left(2^{SR}-1\right)P^{-1}\right)\right\} \\ &\overset{(d)}{\leq} \mathbb{P}\left\{\bigcap_{i=1}^{L}\left(\sum_{j=1}^{l_i}\mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}} + \sum_{j=1}^{l_i-1}\nu_{i,j} \geq 1 - \log\frac{c\left(2^{SR}-1\right)}{P}\right)\right\} \\ &\overset{(e)}{\doteq} \mathbb{P}\left\{\bigcap_{i=1}^{L}\left(\sum_{j=1}^{l_i}\mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}} + \sum_{j=1}^{l_i-1}\nu_{i,j} \geq 1\right)\right\}. \end{aligned} \tag{84}$$

In the above equation, (a) follows from the fact that $1+\lambda_{\max}(\mathbf{A}^{\frac{1}{2}}) \leq |\mathbf{I}+\mathbf{A}|$, for a positive semi-definite matrix $\mathbf{A}$. (b) results from $\mathbf{P}_n \preccurlyeq c\mathbf{I}_L$. (c) follows from the fact that $\lambda_{\max}(\mathbf{H}_T) \geq \max_i \lambda_{\max}(\mathbf{X}_{i,i})$. To obtain (d), we first show that

$$\lambda_{\max}(\mathbf{AUB}) \geq \lambda_{\max}(\mathbf{A})\lambda_{\max}(\mathbf{A})\left|\mathbf{v}_{r,\max}^H(\mathbf{A})\mathbf{U}\mathbf{v}_{l,\max}(\mathbf{B})\right|^2, \tag{85}$$

[12]More precisely, with probability greater than $1-P^{-\delta}$ for any $\delta > 0$.

for any matrices $\mathbf{A}$, $\mathbf{U}$ and $\mathbf{B}$. To show this, we write

$$\begin{aligned}
\lambda_{\max}(\mathbf{AUB}) &= \max_{\substack{\mathbf{x} \\ \|\mathbf{x}\|^2=1}} \left\|\mathbf{x}^H\mathbf{AUB}\right\|^2 \\
&\geq \left\|\mathbf{v}_{l,\max}(\mathbf{A})\mathbf{AUB}\right\|^2 \\
&= \left\|\sigma_{\max}(\mathbf{A})\mathbf{v}_{r,\max}^H(\mathbf{A})\mathbf{U}\sum_i \mathbf{v}_{l,i}(\mathbf{B})\sigma_i(\mathbf{B})\mathbf{v}_{r,i}^H(\mathbf{B})\right\|^2 \\
&\overset{(a)}{=} \sum_i \left\|\sigma_{\max}(\mathbf{A})\mathbf{v}_{r,\max}^H(\mathbf{A})\mathbf{U}\mathbf{v}_{l,i}(\mathbf{B})\sigma_i(\mathbf{B})\mathbf{v}_{r,i}^H(\mathbf{B})\right\|^2 \\
&\geq \left\|\sigma_{\max}(\mathbf{A})\mathbf{v}_{r,\max}^H(\mathbf{A})\mathbf{U}\mathbf{v}_{l,\max}(\mathbf{B})\sigma_{\max}(\mathbf{B})\mathbf{v}_{r,\max}^H(\mathbf{B})\right\|^2 \\
&\overset{(b)}{=} \lambda_{\max}(\mathbf{A})\lambda_{\max}(\mathbf{B})\left|\mathbf{v}_{r,\max}^H(\mathbf{A})\mathbf{U}\mathbf{v}_{l,\max}(\mathbf{B})\right|^2, \qquad (86)
\end{aligned}$$

where $\sigma_i(\mathbf{A})$ denotes the i'th singular value of $\mathbf{A}$, and $\sigma_{\max}(\mathbf{A})$ denotes the singular value of $\mathbf{A}$ with the highest norm. Here, (a) follows from the fact that as $\{\mathbf{v}_{r,i}(\mathbf{B})\}$ are orthogonal vectors, the square-norm of their summation is equal to the summation of their square-norms. (b) results from the fact that $\lambda_i(\mathbf{A}) = |\sigma_i(\mathbf{A})|^2, \forall i$. By recursively applying (85), it follows that

$$\begin{aligned}
\lambda_{\max}(\mathbf{X}_{i,i}) &\geq \lambda_{\max}\left(\mathbf{H}_{K+1,\mathrm{p}_i(l_i-1)}\right)\gamma_{i,l_i-1}\lambda_{\max}\left(\mathbf{H}_{\mathrm{p}_i(l_i-1),\mathrm{p}_i(l_i-2)}\right)\gamma_{i,l_i-2}\cdots\gamma_{i,1}\lambda_{\max}\left(\mathbf{H}_{\mathrm{p}_i(1),0}\right) \\
&= \prod_{j=1}^{l_i}\lambda_{\max}\left(\mathbf{H}_{\mathrm{p}_i(j),\mathrm{p}_i(j-1)}\right)\prod_{j=1}^{l_i-1}\gamma_{i,j}. \qquad (87)
\end{aligned}$$

Noting the definitions of $\mu_{\{i,j\}}$ and $\nu_{i,j}$, (d) easily follows. Finally, (e) results from the fact that as $P \to \infty$, the term $\log \frac{c\left(2^{SR}-1\right)}{P}$ can be ignored.

Since the left and the right unitary matrices resulting from the SVD of an i.i.d. complex Gaussian matrix are independent of its singular value matrix [32] and $\mathbf{U}_{i,j}$ is an independent isotropically distributed unitary matrix, we conclude that all the random variables in the set $\left\{\{\mu_e\}_{e\in E}, \{\nu_{i,j}\}_{1\leq i\leq L, 1\leq j<l_i}\right\}$ are mutually independent. From the probability distribution analysis of the singular values of circularly symmetric Gaussian matrices in [9], we can easily prove $\mathbb{P}\left\{\mu_e \geq \mu_e^0\right\} \doteq P^{-N_aN_b\mu_e^0} = P^{-w_e\mu_e^0}$. Similarly, as $\mathbf{U}_{i,j}$ is isotropically distributed, it can be shown that $\mathbb{P}\left\{\nu(i,j) \geq \nu_0(i,j)\right\} \doteq P^{-\nu_0(i,j)}$. Hence, defining $\boldsymbol{\mu} = [\mu_e]_{e\in E}^T$, $\boldsymbol{\nu} = [\nu_{i,j}]_{1\leq i\leq L, 1\leq j<l_i}^T$, and $\mathbf{w} = [w_e]_{e\in E}$, we have

$$\mathbb{P}\left\{\boldsymbol{\mu} \geq \boldsymbol{\mu}_0, \boldsymbol{\nu} \geq \boldsymbol{\nu}_0\right\} \doteq P^{-(\mathbf{1}\cdot\boldsymbol{\nu}+\mathbf{w}\cdot\boldsymbol{\mu})}. \qquad (88)$$

Let us define $\mathcal{R}$ as the region in $\mathbb{R}^{|E|+\sum_{i=1}^{L} l_i - L}$ of the vectors $\left[\boldsymbol{\mu}^T\boldsymbol{\nu}^T\right]^T$ such that for all $1 \leq i \leq L$, we have $\sum_{j=1}^{l_i}\mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}} + \sum_{j=1}^{l_i-1}\nu_{i,j} \geq 1$. Using the same argument as in the proof of Theorem 3, we conclude that $\mathbb{P}\{\mathcal{R}\} = \mathbb{P}\left\{\mathcal{R}\bigcap\mathbb{R}_+^{|E|+\sum_{i=1}^{L} l_i - L}\right\}$. Hence, defining $\mathcal{R}_+ = \mathcal{R}\bigcap\mathbb{R}_+^{|E|+\sum_{i=1}^{L} l_i - L}$ and $d_0 = \min_{[\boldsymbol{\mu}^T\boldsymbol{\nu}^T]^T\in\mathcal{R}_+} \mathbf{w}\cdot\boldsymbol{\mu} + \mathbf{1}\cdot\boldsymbol{\nu}$, which can easily be verified to be bounded, and applying the same argument as in the proof of Theorem 3, we have

$$\mathbb{P}\{\mathcal{E}\} \dot{\leq} \mathbb{P}\{\mathcal{R}_+\} \doteq P^{-d_0}. \qquad (89)$$

To complete the proof, we have to show that $d_0 = d_G$, or equivalently, $d_0 = L$ (note that $L = d_G$). The value of d_0 is obtained from the following linear programming optimization problem

$$\begin{aligned} & \min \ \mathbf{w} \cdot \boldsymbol{\mu} + \mathbf{1} \cdot \boldsymbol{\nu} \\ & s.t. \quad \boldsymbol{\mu} \geq \mathbf{0}, \boldsymbol{\nu} \geq \mathbf{0}, \forall i \sum_{j=1}^{l_i} \mu_{\{\mathrm{p}_i(j), \mathrm{p}_i(j-1))\}} + \sum_{j=1}^{l_i - 1} \nu_{i,j} \geq 1. \end{aligned} \tag{90}$$

According to the argument of linear programming [33], the solution of the above linear programming problem is equal to the solution of the dual problem which is

$$\begin{aligned} & \max \ \sum_{i=1}^{L} f_i \\ & s.t. \quad \mathbf{0} \leq \mathbf{f} \leq \mathbf{1}, \forall e \in E, \sum_{e \in \mathrm{p}_i} f_i \leq w_e. \end{aligned} \tag{91}$$

Let us consider the solution $\mathbf{f}_0 = \mathbf{1}$ for (91). As the path sequence $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$ consists of the paths that form the maximum flow in $\hat{G}$, we conclude that for every $e \in E$, we have $\sum_{e \in \mathrm{p}_i} 1 \leq w_e$. Hence, $\mathbf{f}_0$ is a feasible solution for (91). On the other hand, as for all feasible solutions $\mathbf{f}$ we have $\mathbf{f} \leq \mathbf{1}$, we conclude that $\mathbf{f}_0$ maximizes (91). Hence, we have

$$d_0 = \min \ \mathbf{w} \cdot \boldsymbol{\mu} + \mathbf{1} \cdot \boldsymbol{\nu} \stackrel{(a)}{=} \max \ \sum_{i=1}^{L} f_i = L = d_G. \tag{92}$$

Here, (a) results from duality of the primal and dual linear programming problems. This completes the proof. ∎

Remark 9- It is worth noting that according to the proof of Theorem 7, any RS scheme achieves the maximum diversity of the wireless multiple-antenna multiple-relays network as long as its corresponding path sequence includes the paths $\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_{d_G}$ used in the proof of Theorem 7.

Theorem 7 shows that the RS scheme is capable of exploiting the maximum achievable diversity gain in multiple-antenna multiple-relay wireless networks. However, as the following example shows, the RS scheme is unable to achieve the maximum multiplexing gain in a general multiple-antenna multiple-node wireless network.

Example- Consider a two-hop relay network consisting of $K = 4$ relay nodes. The transmitter and the receiver are equipped with $N_0 = N_5 = 2$ antennas, while each of the relays has a single receiving/transmitting antenna. There exists no direct link between the transmitter and the receiver, i.e. $\{0, 5\} \notin E$. For the sake of simplicity, assume that the relays are non-interfering, i.e. $1 \leq a \leq 4, 1 \leq b \leq 4, \{a, b\} \notin E$. Let us partition the set of relays into $\mathcal{S}_0 = \{1, 2\}, \mathcal{S}_1 = \{3, 4\}$. Consider the following amplify-and-forward strategy: In the i'th time slot, the relay nodes in $\mathcal{S}_{i \mod 2}$ transmit what they have received in the last time slot, while the relay nodes in $\mathcal{S}_{(i+1) \mod 2}$ receive the transmitter's signal. It can be easily verified

that this scheme achieves a maximum multiplexing gain of $r = 2$. However, the proposed RS scheme achieves a maximum multiplexing gain of $r = 1$.

VI. Conclusion

The setup of a multi-antenna multiple-relay network is studied in this paper. Each pair of nodes are assumed to be either connected through a quasi-static Rayleigh fading channel or disconnected. A new scheme called *random sequential* (RS), based on the amplify-and-forward relaying, is introduced for this setup. It is proved that for the general multiple-antenna multiple-relay networks, the proposed scheme achieves the maximum diversity gain. Furthermore, bounds on the diversity-multiplexing tradeoff (DMT) of the RS scheme are derived for a general single-antenna multiple-relay network. Specifically, 1) the exact DMT of the RS scheme is derived under the assumption of "non-interfering relaying"; 2) a lower-bound is derived on the DMT of the RS scheme (no conditions imposed). Finally, it is shown that for the single-antenna two-hop multiple-access multiple-relay network setup where there is no direct link between the transmitter(s) and the receiver, the RS scheme achieves the optimum diversity-multiplexing tradeoff. However, for the multiple access single relay scenario, we show that the RS scheme is unable to perform optimum in terms of the DMT, while the dynamic decode-and-forward scheme is shown to achieves the optimum DMT for this scenario.

References

[1] Sh. Oveis Gharan, A. Bayesteh, and A. K. Khandani, "Optimum diversity-multiplexing tradeoff in the multiple relays network," University of Waterloo, Tech. Rep. UW-E&CE#2007-19, April 2007, Available in Library and Archives Canada.

[2] ——, "Optimum diversity-multiplexing tradeoff in the multiple relays network," University of Waterloo, Tech. Rep. UW-E&CE#2007-32, September 2007, Available at http://arxiv.org/pdf/0709.4506.

[3] ——, "Optimum diversity-multiplexing tradeoff in the multiple relays network," University of Waterloo, Tech. Rep. UW-E&CE#2007-35, November 2007, Available in Library and Archives Canada.

[4] ——, "On diversity-multiplexing tradeoff in multiple-relay networks," University of Waterloo, Tech. Rep. UW-E&CE#2008-04, March 2008, Available in Library and Archives Canada.

[5] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[6] K. Azarian, H. El Gamal, and Ph. Schniter, "On the Achievable Diversity-Multiplexing Tradeoff in Half-Duplex Cooperative Channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[7] M. Yuksel and E. Erkip, "Cooperative Wireless Systems: A Diversity-Multiplexing Tradeoff Perspective," *IEEE Trans. Inform. Theory*, Aug. 2006, under Review.

[8] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Select. Areas Commun.*, vol. 24, no. 3, pp. 659–672, March 2006.

[9] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073– 1096, May 2003.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New york: Wiley, 1991.

[11] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, “Fading relay channels: Performance limits and space-time signal design,” *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[12] P. Mitran, H. Ochiai, and V. Tarokh, “Space-time diversity enhancements using collaborative communications,” *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 2041–2057, June 2005.

[13] M. Katz and S. Shamai (Shitz), “Transmitting to colocated users in wireless ad hoc and sensor networks,” *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3540–3563, Oct. 2005.

[14] Sh. Yang and J.-C. Belfiore, “Towards the optimal amplify-and-forward cooperative diversity scheme,” *IEEE Trans. Inform. Theory*, vol. 53, pp. 3114–3126, Sept. 2007.

[15] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inf. Theory*, vol. 51, pp. 3037–3063, Sept. 2005.

[16] LL Xie and PR Kumar, “An achievable rate for the multiple-level relay channel,” *IEEE Trans. Inf. Theory*, vol. 51, pp. 1348– 1358, April 2005.

[17] B. Wang, J. Zhang, and A. Host-Madsen, “On the capacity of MIMO relay channels,” *IEEE Trans. Inf. Theory*, vol. 51, pp. 29–43, Jan. 2005.

[18] M. Gastpar, G. Kramer, and P. Gupta, “The multiple-relay channel: Coding and antenna-clustering capacity,” in *IEEE Int. Symp. Information Theory*, 2002, p. 136.

[19] A. S. Avestimehr and D. N. C. Tse, “Outage capacity of the fading relay channel in the low-SNR regime,” *IEEE Trans. Inf. Theory*, vol. 53, pp. 1401–1415, April 2007.

[20] A. S. Avestimehr, S. N. Diggavi and D. N. C. Tse, “A deterministic approach to wireless relay networks,” in *45th Allerton Conference on Communication, Control, and Computing*, 2007.

[21] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *IEEE Trans. Inf. Theory*, vol. 46, p. 388404, March 2000.

[22] L. L. Xie and P. R. Kumar, “A network information theory for wireless communication: scaling laws and optimal operation,” *IEEE Trans. Inf. Theory*, vol. 50, p. 748767, May 2004.

[23] M. Gastpar and M. Vetterli, “On the capacity of large Gaussian relay networks,” *IEEE Trans. Inf. Theory*, vol. 51, pp. 765–779, March 2005.

[24] H. Bolcskei, R. U. Nabar, O. Oyman, and A.J. Paulraj, “Capacity Scaling Laws in MIMO Relay Networks,” *IEEE Trans. on Wireless Communications*, vol. 5, no. 6, pp. 1433–1444, June 2006.

[25] Sh. Oveis Gharan, A. Bayesteh, and A. K. Khandani, “Asymptotic analysis of amplify and forward relaying in a parallel MIMO relay network,” *IEEE Trans. Inf. Theory*, 2006, revised, available online at http://cst.uwaterloo.ca/.

[26] VI Morgenshtern, H Bolcskei, RU Nabar, “Distributed orthogonalization in large interference relay networks,” in *IEEE Int. Symp. Information Theory*, 2005, pp. 1211–1215.

[27] V. I. Morgenshtern and H. Bolcskei, “Crystallization in large wireless networks,” *IEEE Trans. Inf. Theory*, vol. 53, pp. 3319–3349, Oct. 2007.

[28] K. Sreeram, S. Birenjith and P. Vijay Kumar, “Multi-hop cooperative wireless networks: diversity multiplexing tradeoff and optimal code design,” February 2008, Available online at http://arxiv.org/pdf/0802.1888.

[29] R. Diestel, *Graph Theory*. Springer, 2006.

[30] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[31] T. M. Cover and A. El Gamal, “Capacity Theorems for the Relay Channel,” *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.

[32] A. M. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*. Now Publishers Inc, 2004.

[33] A. Schrijver, *Theory of Linear and Integer Programming*. Wiley, 1998.